Sonar Image Based Advanced Feature Extraction and Feature Mapping Algorithm for Under-Ice AUV Navigation

Herath Mudiyanselage Dinuka Doupadi Bandara,
National Centre for Maritime Engineering and Hydrodynamics
Australian Maritime College

Submitted in partial fulfilment of the requirements for the degree of Master of Philosophy

University of Tasmania

July 2017

[Page intentionally left blank]

DECLARATIONS

Declaration of Originality and Authority of Access

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

This thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968.

Herath Mudiyanselage Dinuka Doupadi Bandara (July 2017)

ACKNOWLEDGEMENT

It is hard to believe my wonderful journey is coming to an end. However, the completion of this thesis marks the beginning of another great adventure ahead. Thus, it is with great pleasure that I thank the many people who made this thesis possible.

First and foremost I would like to thank my supervisors Dr Hung Nguyen, Assistant Professor Alexander Forrest, Dr Shantha Jayasinghe and Dr Zhi Quan Leong for their invaluable guidance, consistent support, good company and encouragement throughout my research work. It is difficult to overstate my gratitude to them as they have not only been great supervisors but also great mentors.

I would also wish to express my gratitude to Mr Peter King for his valuable support and guidance in this research. Also, I would like to thank the AUV team at AMC including Dr Damien Guihen and Dr Konrad Zurcher for their support and advice regarding my research. Moreover, I would like to thank all my friends who have created a very unforgettable memory in this journey as well as staff members at AMC and UTAS.

Sonar images were provided by the Responsive AUV Localisation and Mapping (REALM) project, supported by the Atlantic Canada Opportunities Agency Atlantic Innovation Fund, Research & Development Corporation Newfoundland and Labrador, Fugro GeoSurvey's Inc. and Memorial University of Newfoundland.

Lastly, and most importantly, I wish to thank my beloved parents for their countless efforts and continuous encouragement throughout my academic career and most special thanks go to my husband, who is giving support to fulfil my dreams.

[Page intentionally left blank]

ABSTRACT

Navigation and localisation of AUVs are challenging in underwater or under-ice environments due to the unavoidable growth of navigational drift in inertial navigation systems and Doppler velocity logs, especially in long-range under-ice missions where surfacing is not possible. Similarly, acoustic transponders are time consuming and difficult to deploy. Terrain Relative Navigation (TRN) and Simultaneous Localisation and Mapping (SLAM) based technologies are emerging as promising solutions as they require neither deploying sensors nor the calculation of distance travelled from a reference point in order to determine the location. These techniques require robust detection of the features present in sonar images and matching them with known images. The key challenge of under-ice image based localisation comes from the unstructured nature of the seabed terrain and lack of significant features. This issue has motivated the research project presented in this thesis. The research has developed technologies for the robust detection and matching of the available features in such environments.

In aiming to address this issue, there are number of feature detectors and descriptors that have been proposed in the literature. These include Harris corner, Scale-Invariant Feature Transform (SIFT), Binary Robust Invariant Scalable Keypoints (BRISK), Speeded-Up Robust Features (SURF), Smallest Univalue Segment Assimilating Nucleus (SUSAN), Features from Accelerated Segment Test (FAST), Binary Robust Independent Elementary Features (BRIEF) and Fast Retina Keypoints (FREAK). While these methods work well in land and aerial complex environments, their application in under-ice environments have not been well explored. Therefore, this research has investigated the possibility of using these detector and descriptor algorithms in underwater environments. According to the test carried out with side-scan sonar images, the SURF and Harris algorithms were found to be better in repeatability while the FAST algorithm was found to be the fastest in feature detection. The Harris algorithm was the best for localisation accuracy. BRISK shows better immunity noise compared to BRIEF. SURF, BRISK and FREAK are the best in terms of robustness. These detector and descriptor algorithms are used for a wide range of varying substrates in underwater environments such as clutter, mud, sand, stones, lack of features and effects on the sonar images such as scaling, rotation and non-uniform intensity and backscatter with

filtering effect. This thesis presents a comprehensive performance comparison of feature detection and matching of the SURF, Harris, FAST, BRISK and FREAK algorithms, with filtering effects. However, these detectors and descriptors have reduced efficiency in underwater environments lacking features. Therefore, this research further addressed this problem by developing new advanced algorithms named the 'SURF-Harris' algorithm, which combined Harris interest points with the SURF descriptor, and the 'SURF-Harris-SURF' algorithm which combined Harris and SURF interest points with the SURF descriptor, using the most significant factor of each detector and descriptor to give better performance, especially in feature mapping.

The major conclusion drawn from this research is that the 'SURF-Harris-SURF' algorithm outperforms all the other methods in feature matching with filtering even in the presence of scaling and rotation differences in image intensity. The results of this research have proved that new algorithms perform well in comparison to the conventional feature detectors and descriptors such as SURF, Harris, FAST, BRISK and FREAK. Furthermore, SURF works well in all the disciplines with higher percentage matching even though it produces fewer keypoints, thus demonstrating its robustness among all the conventional detectors and descriptors. Even if there are a large number of features in a cluttered environment, it produces less matching compared to features that are distributed. Another conclusion to be derived from these results is that the feature detection and matching algorithms performed well in environments where features are clearly separated. Based on these findings, the comprehensive performance of combined feature detector and descriptor is discussed in the thesis. This is a novel contribution in underwater environments with sonar images. Moreover, this thesis outlines the importance of having a new advanced feature detector and mapping algorithm especially for sonar images to work in underwater or under-ice environments.

[Page intentionally left blank]

Contents

List of Figures

List of Tables

List of Abbreviations

AUG	Autonomous Glider
AUV	Autonomous Underwater Vehicle
BRIEF	Binary Robust Independent Elementary Feature
BRISK	Binary Robust Invariant Scalable Keypoints
CML	Concurrent Mapping and Localisation
DoH	Determinant of Hessian
DT	Distance Travel
DVL	Doppler Velocity Log
EKF	Extended Kalman Filter
FAST	Features from Accelerated Segment Test
FREAK	Fast Retina Keypoints
GIB	Global Positioning System Intelligent Buoys
GPS	Global Positioning System
INS	Inertial Navigation Systems
LBL	Long Baseline
LOG	Laplacian of Gaussian
MBES	Multi-Beam Echo Sounder
MSAC	M-Estimator Sample Consensus
PSNR	Peak Signal to Noise Ratio
RANSAC	Random Sample Consequence
SIFT	Scale-Invariant Feature Transform
SLAM	Simultaneous Localisation and Mapping

SSD	Sum of Squared Difference
SSS	Side-Scan Sonars
SURF	Speeded-Up Robust Features
SUSAN	Smallest Univalue Segment Assimilating Nucleus
TRN	Terrain Relative Navigation
USBL	Ultra-Short Base Line

Nomenclature

(x, y)	Given point of the image I
I_x	Partial derivatives of x in Harris detector
I_y	Partial derivatives of y in Harris detector
$L_{xy}(\mathbf{x}, \sigma)$	Convolution of the Gaussian second order derivative with image I in point *xy*
$L_{yy}(\mathbf{x}, \sigma)$	Convolution of the Gaussian second order derivative with image I in point *y*
V_{DVL}	Velocity of the DVL in the navigation frame
V_c	Water current velocity in the navigation frame
d_x, d_y	Derivatives in the x and y for sub-regions
$r_{b,1}, r_{b,2}, r_{b,3}\ and\ r_{b,4}$	Unit vector along the beam
v_1,v_2, $v_3\ and\ v_4$	Four beams, assumed to be mounted at 30^0 degrees from the vertical in the DVL
(x, y)	Difference in intensity for a displacement in all direction
λ1, λ2	Eigenvalues of Matrix C
$E(x, y)$	Corner detection
$H(\mathbf{x}, \sigma)$	Hessian matrix
I	Image patch over the area of $(u, v)$and shifted by (x, y)
$I(u, v)]^2$	Intensities
$I(x + u, y + v)$	Shifted intensities

$L_{xx}(\mathbf{x}, \sigma)$	Convolution of the Gaussian second order derivative with image I in point
n	Layers of the pyramid
Ci	Octaves of the pyramid
di	Intra-octaves of the pyramid
P	Total matrix of Harris features and Hessian features in matrix form
$P1$	Calculated Hessian features (matrix form)
$P2$	Calculated Harris features (matrix form)
V	Haar wavelent response vector for horizontal and vertical (Sub-region)
$c(x, y)$	Intensity structure of the local neighbourhood
$w\ (u, v)$	Window function
σ	Scale of the image I

[Page intentionally left blank]

Chapter 1 Introduction

1.1 Background

Autonomous underwater vehicles (AUVs) are becoming increasingly popular in underwater operations such as search and rescue (Jacoff et al. 2015), mapping (Caress et al. 2012), climate change assessments (Schofield et al. 2010), marine habitat monitoring (Williams et al. 2012), shallow water mine countermeasures (Freitag et al. 2005), pollutant monitoring (McCarthy & Sabol 2000) and under-ice inspection (McFarland et al. 2015). As the name suggests, AUVs should be able to determine their location and course without human intervention. In order to achieve this, AUVs need to address two critical problems: 1) mapping the environment through which they are transiting and 2) finding their location relative to the map. Although GPS is a widely used positioning method in land and aerial vehicles, AUVs cannot rely on it underwater due to the rapid attenuation of radio signals in water (Paull et al. 2014). In addition, current underwater localisation systems tend to rely on baseline transponders. However, these techniques are not able to provide accurate position estimates for long-range missions, especially for AUVs operating in under-ice environments.

Over the last few decades, AUV operations have continued to push the limits of the existing technology and return measurements that would be logistically impossible by other means. Nevertheless, fully autonomous navigation under ice still requires further significant development since:

- Under-ice environments are largely low contrast compared to land terrain, often featureless (e.g. smooth ice) and are comprised mostly of repetitive ice textures (Spears et al. 2015);
- Navigation under translating and rotating surfaces such as icebergs remains an unsolved problem (McFarland et al. 2015);
- Unavoidable growth of navigational drift in the position estimates (Webster et al. 2015); and
- Requirement of more time and difficulty in deployment of transponders (Medagoda et al. 2016).

There are various potential AUV internal and external sensors, and state estimators to aid navigation in under-ice environments. Internal sensors such as the Inertial Navigation System (INS) do not measure position, but rather determine the location internally by integrating instantaneous vehicle velocities or accelerations. External systems determine positions relative to the properties or features of the environment, and state estimators represent the algorithms used for underwater localisation and mapping such as Terrain Relative Navigation (TRN) and Simultaneous Localisation and Mapping (SLAM). Sonar imaging, Doppler Velocity Log (DVL), and underwater acoustic positioning systems are some of the sensor-based navigation support systems. Out of these solutions, sonar imaging is the most feasible method as it uses natural features present in the environment (Kimball & Rock 2011; Richmond et al. 2009; Miller et al. 2010). Apart from that, sonar imaging is free from the drift and thus no extra recalibration effort is required. Therefore, in recent years, there has been an increased research interest in sonar imaging based AUV localisation in academia as well as in the industry (Bandara et al. 2016; Li et al. 2017; Song et al. 2016).

Sonar imaging based localisation has three main elements, namely: feature detection, feature description and feature matching. In feature detection, key features present in the sonar image are identified. In order to achieve fast localisation, the detector should be able to capture prominent features and avoid features or patterns that are common to most of the areas surveyed. At the same time the feature detection method should be at the point of reliability not to miss important features. The identified features may vary from one to the other, having different shapes and sizes. Therefore, each identified feature has to be uniquely represented in terms of its pixel distribution, which is done using the feature descriptor. The third element, the feature-matching algorithm, uses these descriptions to match with similar descriptions found in stored maps (Fraundorfer & Scaramuzza 2012; King et al. 2013; Vandrish et al. 2011; Zerr et al. 2005). As the stored images and the current images may have different scales and orientations, the descriptor should be made immune to scale and rotation, which makes developing a suitable descriptor a challenging task (Fraundorfer & Scaramuzza 2012). Therefore, success of sonar image processing based localisation heavily depends on the performance of its detector and descriptor.

1.2 Problem Definition

Extracting reliable key features in dynamic and unstructured underwater environments is the major challenge in underwater navigation (Chen & Hu 2011; Guth et al. 2014). This becomes even more difficult when dealing with the unstructured nature of ice terrains which are largely low contrast, featureless and comprised mostly of repetitive ice texture (Spears et al. 2015) as shown in Figure 1-1. Moreover, if the entire surveyed region lacks texture and variations, determining the accurate location of the AUV is virtually impossible without advanced feature detection methods.

There are a number of feature detectors and descriptors reported in the literature that have been developed primarily for land and aerial vehicles. Among these, the most popular feature detectors and descriptors are Harris corner (Harris & Stephens 1988) detectors, Smallest Univalue Segment Assimilating Nucleus (SUSAN) (Smith 1992), Scale-Invariant Feature Transform (SIFT) (Lowe 1999), Speeded-Up Robust Features (SURF) (Bay et al. 2006), Features from Accelerated Segment Test (FAST), (Rosten & Drummond 2006), Binary Robust Independent Elementary Features (BRIEF) (Calonder et al. 2010), Binary Robust Invariant Scalable Keypoints (BRISK) (Leutenegger et al. 2011) and Fast Retina Keypoints (FREAK) (Alahi et al. 2012).

While these feature detectors and descriptors work well in land and aerial complex environments their efficiency is low in underwater environments that lack significant features. Furthermore, their applications in under-ice environments have not been well explored. Therefore, this thesis has investigated the possibility of using new advanced feature detection algorithms in underwater environments, especially in low contrast and featureless environments. Furthermore, this proposed new algorithm should have the potential to incorporate existing localisation techniques such as SLAM and TRN.

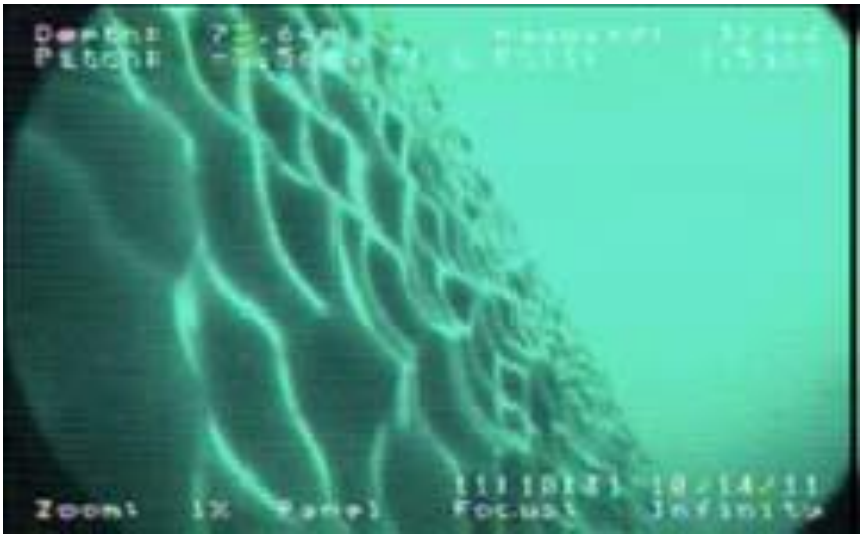


Figure 1-1. Under-ice terrain (low contrast and repetitive ice-texture) on the side wall of the Petermann Ice Island (Forrest et al. 2012)

[This figure is used with permission from author, Alexander L Forrest "Digital terrain mapping of Petermann Ice Island fragments in the Canadian high arctic. 21st IAHR International Symposium on Ice, 2012. 1-12]

1.3 Scope of the Thesis

The motivation behind this study is to develop new advanced algorithms for feature detection and feature matching to work in underwater environments which are lacking in significant features or variations. These environments are largely low contrast compared to land terrain and are mostly featureless, as shown in the side-scan sonar image of the seabed in Figure 1-2. Therefore, the specific research question for this thesis is:

'Can a sonar image based feature detector and descriptor algorithm be developed to provide more reliable results for AUV navigation in under-ice environments which lack significant texture and distinguishable patterns?'

Moreover, this thesis outlines the importance of developing an advanced feature detection and mapping algorithm, especially for sonar images to work in underwater or under-ice environments.

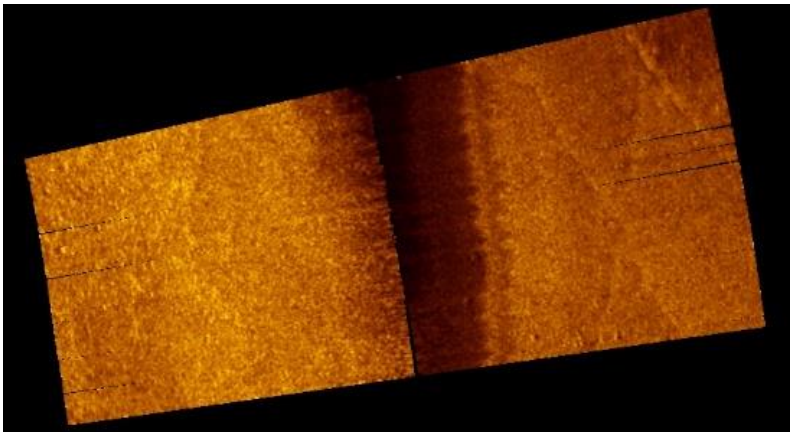


Figure 1-2. Side-scan sonar image with lack of significant features

1.3.1 Objectives

While answering the above-mentioned research question of the thesis, the following objectives were achieved:

- Identified suitable keypoint detectors and descriptors that can be used for side-scan sonar images based localisation and mapping.
- Developed two novel keypoint detector and matching algorithms to work in underwater environments.

1.4 Expected Outcomes

- To identify potential keypoint detectors and descriptors to work with side-scan sonar images in underwater environments and assess their performance with regard to filtering effect.
- To develop a new advanced algorithm for extracting and mapping key features in underwater environments that lack features.

1.5 Principal Contributions

The principal contributions of this thesis include:

- The development and testing of two new advanced algorithms named ‘SURF-Harris’, which combines Harris interest points with SURF descriptor, and ‘SURF-Harris-SURF’ which combines Harris and SURF interest points with SURF descriptor to yield better performance, especially in feature mapping; and

- Investigation into the effects of filtering applied to conventional feature detection and description in side-scan sonar images collected over two consecutive days, covering a wide range of scenarios that can occur in an underwater environment.

1.6 Methodology

To achieve the outcomes of this thesis, the research questions are addressed through three main components:

- A review of the literature on technologies for underwater and under-ice AUV navigation and localisation, especially those using side-scan sonar images.
- Analysis of potential keypoint detector and descriptor algorithms which can be used for side-scan sonar images in a wide range of scenarios that can occur in underwater environments such as clutter, mud, sand, stones and lack of features with filtering effect.
- Develop new advanced feature detection algorithms to work in environments that lack features and validate these through simulations conducted in the MATLAB software environment.

1.7 Thesis Structure

Chapter 1: Introduces the background, research questions, objectives and the direction of the research on 'developing a sonar image based feature detector and descriptor algorithm for AUV navigation'. Major contributions and the organisation of the thesis are also summarised in this introductory chapter.

Chapter 2: Outlines and discusses the literature surrounding underwater AUV navigation and localisation. The direction for subsequent work is framed to incorporate existing work on the subject, while identifying areas of further work that are possible.

Chapter 3: Starts with evaluating potential feature (keypoint) detectors and descriptors which can be used for side-scan sonar images in underwater environments. According to the tests carried out with side-scan sonar images, SURF, Harris, FAST, BRISK and FREAK feature detectors and descriptors are selected. These algorithms are used for a wide range of scenarios that can occur in underwater environments such as clutter, mud, sand, stones and lack of features, and effects on the sonar images such as scaling, rotation and non-uniform intensity and backscatter with filtering effect. Relevant simulation results and major problems associated with the conventional detectors and descriptors are explained in detail.

Chapter 4: Outlines and discusses a novel contribution of new advanced algorithms 'SURF-Harris' and 'SURF-Harris-SURF' which give better performance, especially in feature mapping. Relevant methodology and simulation results are explained in detail.

Chapter 5: Presents a summary of the thesis, conclusions drawn from the study and suggests avenues for future research.

Chapter 2 Literature Review

Autonomous platforms have been used under ice since the early 1970s (Francois et al. 1972). Since then, AUVs have been used for numerous underwater operations and under-ice inspections. These under-ice missions have pushed the limits of existing technology and provided highly valuable scientific information, but the ability of an autonomous platform to consistently estimate its own geo-referenced position in real time with acceptable accuracy continues to be a challenge. Current underwater localisation systems tend to rely on baseline transponders and are unable to provide accurate position estimates for long-range missions. This is especially true for AUVs operating in under-ice environments. This chapter provides a comprehensive review of suitable sensors and navigation methods for under-ice AUV navigation followed by new methodologies proposed in subsequent chapters, which can be implemented for future developments for under-ice navigation, especially in featureless environments.

As mentioned in Chapter 1, AUV navigation and localisation techniques in under-ice environments can be categorised as sensors and state estimators, as illustrated in Figure 2-1. Inertial Navigation System (INS) and Doppler Velocity Log (DVL) navigation belong to internal sensors and sonar imaging and underwater acoustic positioning comes under the roof of external sensors. Furthermore, state estimators represent the algorithms used for underwater localisation and mapping such as Terrain Relative Navigation (TRN) and Simultaneous Localisation and Mapping (SLAM). Although TRN has been studied for many years, development for underwater missions requires real-time accurate terrain data collection from available sensors and high confidence matching algorithms. As this still has not been developed, TRN has yet to become a practical solution for underwater navigation. Limited work has been conducted using underwater SLAM for under-ice navigation but has yet to see broad application (Stone Aerospace/PSC, 2016).

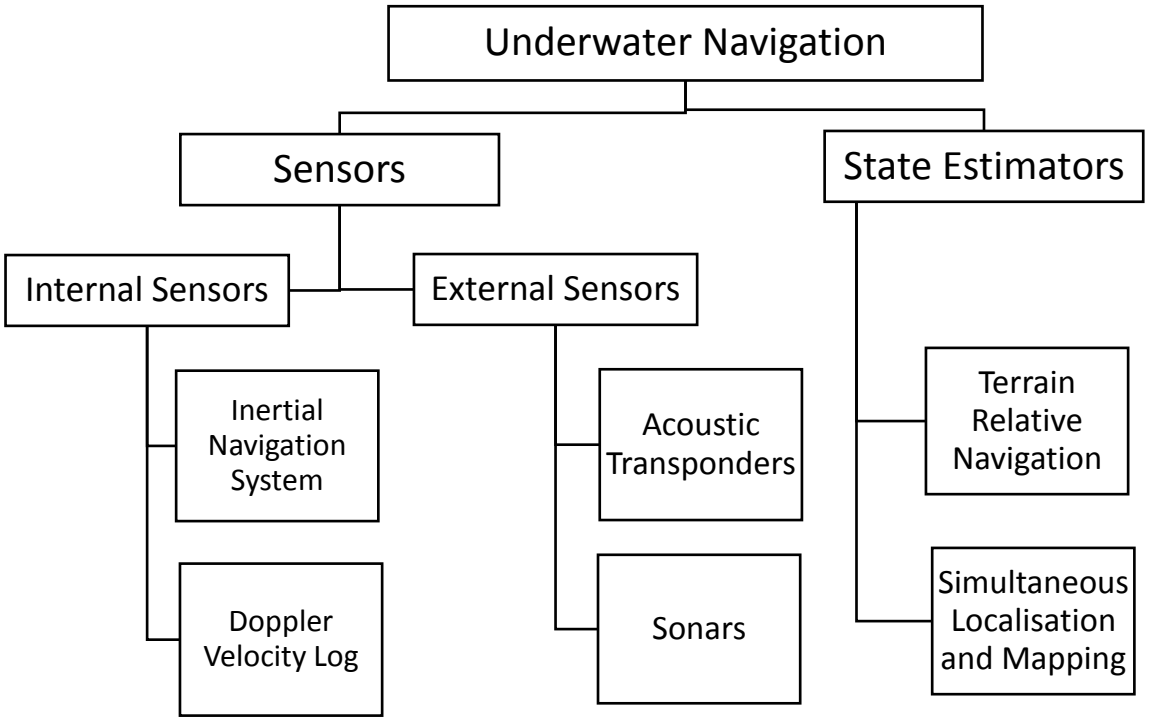


Figure 2-1. Classification of AUV navigation and localisation techniques

2.1 Internal Sensors

There are various sensors deployed in AUVs for under-ice navigation and localisation such as Inertial Navigation Systems (INSs) and Doppler Velocity Logs (DVLs). INS units navigate relative to the initial position. They require accurate knowledge of the vehicle state which depends on sensors to provide measurements of the derivatives of the states (Hildebrandt et al. 2013).

2.1.1 Inertial Navigation Systems (INS)

INS units are an advanced form of dead reckoning that consists of an embedded computer, motion sensors (accelerometer), and rotation sensors (gyroscope) to continuously calculate the position, orientation and velocity of a moving object without the need for external references. According to Grewal et al. (2007) the main advantages of INS over other forms of navigation are autonomy and independency from other external aids. In addition, an INS is suitable for integrated navigation, guidance and control. Most of the time, an INS corrects its position measurements using a GPS receiver. A GPS/INS integrated system can achieve relatively high accuracy of localisation in a shallow water environment within a short

period of time with regular surfacing for GPS fixes. However, INS accumulated errors have effects on localisation error if they are not corrected by GPS for a long time (Chen et al. 2013). Therefore, GPS/INS is not suitable for under-ice long-range missions where regular surfacing for GPS fixes is not possible. Kinsey et al. (2006) showed navigation-grade INS with 0.01^0 gyro bias. This bias translates into a ~1 km/h position drift without aiding. Therefore, INS is a better solution for navigation systems for both sea floor and ice surface with DVL. However, one of the biggest challenges of using these units in polar regions is to achieve alignment at high latitude where rotational accelerations are at a minimum.

2.1.2 Doppler Velocity Log (DVL)

A Doppler Velocity Log (DVL) uses acoustic measurements to capture bottom tracking and determines the velocity vector of an AUV moving across the seabed, as shown in Figure 2-2. It determines the AUV surge, sway and heave velocities by transmitting acoustic pulses and measuring the Doppler shifted returns from the pulses off the seabed (Paull et al. 2014).

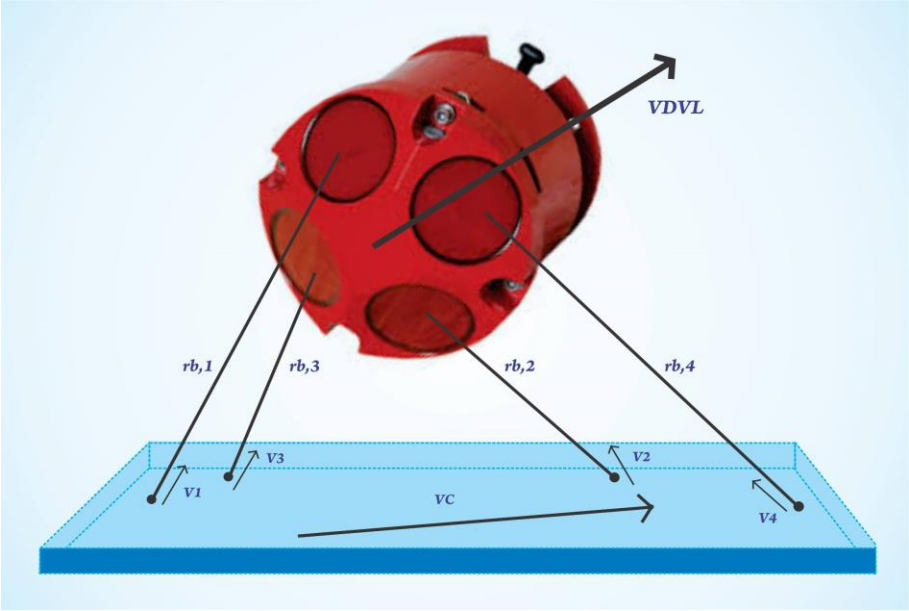


Figure 2-2. Operation of the DVL sensor in water track mode (Medagoda et al. 2016)

V_{DVL} - Velocity of the DVL in the navigation frame

V_c - Water current velocity in the navigation frame

v_1,v_2, v_3 and v_4 - Four beams, assumed to be mounted at 30^0 from the vertical in the DVL

$r_{b,1}$,$r_{b,2}$, $r_{b,3}$ $and r_{b,4}$ - Unit vector along the beam

According to the previous field experiment reports, the biggest challenge in DVL-based navigation is the drift in the position estimates (Kaminski et al. 2010; McEwen et al. 2005; Richmond et al. 2009). This becomes even more difficult in long-range AUV navigation (Medagoda et al. 2016). Nevertheless, a DVL is hardly ever used alone for underwater navigation and is generally combined with other sensors such as acoustic transponders and INS units (Hou et al. 2008; Lin & Wei 2004; Rigby et al. 2006). Therefore, it is important to use another feasible methodology such as combination of TRN and SLAM with sonars for under-ice missions rather than conventional sensors. Moreover, Kimball & Rock (2011) revealed that existing navigation methods which rely on internal sensors are not adequate to enable navigation with respect to free-floating icebergs. Presented in their paper (Kimball & Rock 2011) was an alternative approach that extended TRN techniques to deal explicitly with iceberg motion using sonar data.

In addition to the examples cited above, the British Columbia Environmental Fluid Mechanics group deployed a *Gavia* AUV in Pavilion Lake, British Columbia and Ellesmere Island, Nunavut in 2008. The selection of these sites was based on finding a location where a distinct delineation existed between first-year and multiyear ice. The *Gavia* AUV was equipped with a GeoSwath™ sonar unit, INS and RDI 1200 kHz DVL. For sub-surface measurements, acoustic long baseline (LBL) or DVL have been used. This operation successfully demonstrated under-ice field trials in both lake-ice and sea-ice capability with a small AUV. Furthermore, it showed refinement of the integration of an INS/DVL system with a LBL system (Forrest et al. 2008).

2.2 External Sensors

2.2.1 Acoustic Transponders

Acoustic transponders measure positions relative to a framework of baseline stations as shown in Figure 2-3, which must be deployed prior to operations. They are generally categorised into the following two types:

- Long Baseline System (LBL); and
- Ultra-Short Baseline System (USBL).

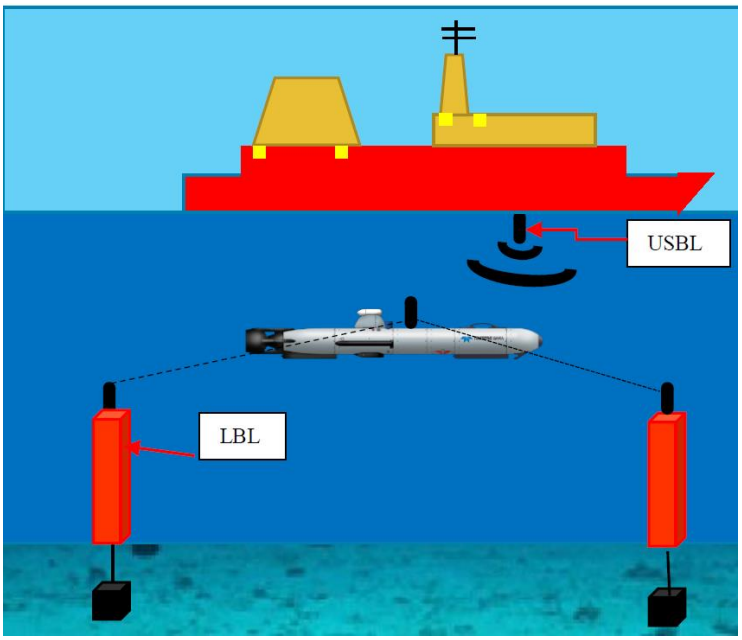


Figure 2-3. Communication networks

Long Baseline (LBL) systems use a sea-flow baseline transponder network and derive the position with respect to the network. As per Figure 2-3, LBL transponders are typically mounted in the corners of the operation site. The position is generated from adding three or more time-of-flight ranges to the sea flow station using triangulation. However, due to the rapid attenuation of high-frequency sound in water, the LBL system typically has a very limited maximum range (Kinsey et al. 2006). On the other hand, Ultra-Short Baseline (USBL) systems rely on a surface transponder, as shown in Figure 2-3. USBL calculates subsea positions by calculating multiple distances, applying triangulations and phase differencing through an array of transceivers (Chen et al. 2013; Paull et al. 2014). This system does not require a

seafloor-mounted system as it uses the ship's hull. The major disadvantage is that its positioning accuracy is dependent on the size of the baseline (Paull et al. 2014).

A few authors have discussed the acoustic transponder design and performance of AUV navigation systems for under-ice environments. In the mid-1980s, Vestgard suggested an under-ice Long Baseline (LBL) positioning system based on a combination of ice-moored and seafloor beacons (Vestgård 1985). This work additionally presented the methodology for a self-calibrating array of acoustic hydrophones and pingers affixed to moving ice-floes that was adapted for AUV monitoring (Jakuba et al. 2008). The most commonly used LBL navigation can be seen in vehicles such as Odyssey (Deffenbaugh et al. 1993). Similarly, Ultra-Short Base Line (USBL) systems were implemented on vehicles such as Remus (Kukulya et al. 2010). The advantage of the USBL is that it does not require a seafloor-mounted system. The major disadvantage is that its positioning accuracy is not as good as the LBL system.

In 2015, the real-time under-ice acoustic navigation system was introduced and the acoustic navigation beacons provided reliable acoustic range estimates and data transfer to an Autonomous Underwater Glider (AUG) through the water column out to 100 km with approximately 50% throughput (Webster et al. 2015). Several operational difficulties were faced during the field experiments such as coordinating acoustic transmissions and although the AUG residence in the sound channel should improve throughput, the data packet was not successfully decoded in the post-processing stage and there was a clock problem. Ensuring that these platforms consistently estimate their own geo-referenced position in real time with acceptable accuracy remains a challenge.

2.2.2 Side-Scan Sonar and Multi-Beam Sonar

Sonar is a widely used type of range sensor in AUV applications. Sonar sensors are based on acoustic signals. There are two main types of sonar used in underwater vehicles: side-scan sonars (SSS) and multi-beam sonars (Chen et al. 2013). A side-scan sonar operates by emitting a single pulse of acoustic energy into the water column and then receiving the reflection signal (echo) (Padial et al. 2013). The main advantage of side-scan sonars is that they can work at relatively high speed to give high area coverage. Unfortunately, image

resolution is inversely proportional to the range (Paull et al. 2014). Multi-beam sonars form separate beams with narrow azimuthal and wide elevation beam-widths using a phased array of transducers (Padial et al. 2014) Multi-beam sonars are capable of gathering echo sounding data more efficiently than single-beam sonars. Nevertheless, the image resolution is inversely proportional to frequency.

Side-scan sonar (SSS) imaging is a promising technique to overcome the challenge caused by the unavoidable growth of drift in vehicle position estimation in underwater vehicle navigation and localisation as it uses features in the environment that are naturally present in order to determine the location. Nevertheless, the particular challenge of under-ice sonar image based localisation can come from the unstructured nature of the terrain with largely low contrast and lack of features. Thus, extracting key features in such an environment is the major challenge. These detectors and descriptors work well in land and aerial complex environments, however they have reduced efficiency in featureless environments such as under ice. Consequently, AUV navigation in under-ice terrain is virtually impossible without advanced feature extracting algorithms and a combination of detectors and descriptors has to be used for under-ice environments.

Each navigation system in the above-described deployments estimated vehicle position in the inertial space. As an example, the dead-reckoned solutions used initially referenced accelerations and velocities, and incorporated drift corrections defined in inertial space. Some deployments corrected for drift using GPS updates at the surface while others used acoustic arrays where positions were fixed and surveyed in inertial space. Hence, the unavoidable growth of navigational drift in the position estimate of submerged vehicles is a persistent issue in underwater robotics. Therefore, sonar imaging is a promising technique with which to overcome the above-mentioned challenges as it is based on detection, identification and classification of features in the environment. Thus, extracting key features in such an environment is the major challenge.

2.3 State Estimators for Under-Ice Localisation and Mapping

2.3.1 Terrain Relative Navigation (TRN)

TRN uses real-time sensing and a terrain or landmark map to determine a vehicle position. TRN was initially developed for aerial platforms and first deployed as the cruise missile guidance system, i.e. TERCOM (Terrain Contour Matching), which performed batch correlation of altimeter measurements against a prior map to generate an estimation of current position (Golden 1980). TRN can eliminate the need for underwater vehicles to surface repeatedly for GPS fixes or work with pre-surveyed acoustic transponder arrays. Multiple research publications have addressed this problem by using prior available maps (Blackinton et al. 1983; Singh et al. 1995), where prior maps were not available but were constructed incrementally using sensor data, and where navigation was achieved by using altitude control and obstacle avoidance without explicit maps (Di Massa & Stewart Jr 1997).

Kimball & Rock (2011) proposed a new iceberg relative navigation technique for mapping and localisation. In the mapping process, the AUV circumnavigates an iceberg by collecting multi-beam sonar ranges to the iceberg's submerged surface that are then combined in post-processing to form a self-consistent iceberg map. The post-processing operation uses terrain correlation at trajectory self-intersections to account for iceberg motion during data collection. In the localisation step the vehicle uses the iceberg map in an extended TRN-based localisation estimator to determine its position with respect to the iceberg. The localisation estimator has been augmented with states that enable it to account for iceberg motion between sonar measurements.

The major challenge in TRN is that the majority of experiments use a sensor combination of DVL/INS and due to the noise generated in these sensor measurements, the vehicle position estimate will rapidly grow without bounds. To overcome this issue, there are possible absolute positioning systems that can be used such as LBL, USBL and GPS Intelligent Buoys (GIB). Nevertheless, these systems require more time and present difficulties for deployment. Therefore, the Multi-Beam Echo Sounder (MBES) is a good solution to overcome this problem as it is able to establish bathymetric maps due to its high resolution data,

capturing a high number of sound points instantly and yielding rich information (Chen et al. 2015).

Although TRN has been studied for under-ice missions (Kinsey et al. 2006), development in real-time accurate terrain data collection from available sensors and a high confidence matching algorithm are still required for TRN to become a practical solution.

2.3.2 Simultaneous Localisation and Mapping (SLAM)

Simultaneous Localisation and Mapping (SLAM), sometimes called Concurrent Mapping and Localisation (CML), is a technique used by autonomous technologies to build up a map with an unknown environment or to update a map of a known environment. Over the last two decades, many research studies were carried out to solve the SLAM problem with various algorithms for land and surface vehicles. However, underwater SLAM has many more challenging issues compared to land SLAM, due to the unstructured nature of the underwater scenarios and difficulty in identifying reliable features.

SLAM can be performed online or offline. Where the current pose is estimated along with the map is called online SLAM (Paull et al. 2014). Most of the authors have developed filters in information form to address the online SLAM. The Bayesian filter is the fundamental filter derived other filters such as Particle filters, Kalman filters and Information filters. Extended Kalman Filter (EKF) is the most popular solution among the SLAM algorithm due to its simplicity. Most of the under-water navigation missions used offline SLAM due to difficulty in pre-processing. Roman et al. (Roman & Singh 2005) showed that the vehicle trajectory and map estimates are using offline SLAM. Furthermore, the estimation of ice-berg trajectory too was performed in 2015, using offline SLAM (Kimball & Rock 2015). There are two key advantages in the computation of an iceberg map offline as a post-processing over an online SLAM formulation. First, it affords the opportunity for a human operator to clean the mapping sonar data and provides input as to which sonar soundings overlap for terrain correlation. Second, it allows a large number of computationally intensive terrain correlation computations to be carried out in machines that are far more powerful than those available on an AUV (Kimball & Rock 2011).

Kimball and Rock's (Kimball & Rock 2011) proposed mapping technique for icebergs was an offline SLAM. It involved circumnavigation of an iceberg using a simple wall-following robotic behaviour, during which the mapping sonar data and initially referenced vehicle navigation data were simply recorded for later use. Fairfield et al. (2007) deployed a full three-dimensional SLAM capability for the DEPTHX vehicle which used a steradian pencil-beam sonar used for operations in flooded sinkholes and tunnels. Furthermore, the team of Stone

Aerospace deployed ENDURANCE AUV, a redesign of the DEPTHX vehicle in West Lake Bonney in Antarctica (Stone Aerospace/PSC 2016).

One of the key challenges to implementing SLAM in a real underwater environment is the limited accuracy of the sensors, particularly for the environments with low light, strong ocean currents and turbid waters. Apart from that, filtering the noise in sonar images and extracting important features in the image are important tasks in SLAM (Guth et al. 2014). Extracting good features of the environment is a highly complex problem, especially when it comes to underwater environments. This problem is quite pronounced when dealing with flat terrain or a poor feature extraction environment. Many underwater features are scale dependent, i.e. sensitive to viewing angle and scale (Chen et al. 2013). The computational complexity of SLAM applications are closely related to the size of the environment to be explored and the methods used for feature extraction, tracking, data association and filtering (Guth et al. 2014).

Image-based localisation is a relatively well-studied field in land and aerial environments. Nonetheless, its application to the underwater domain is in its infancy. However, most of the relevant work on AUVs utilises sonar imaging as part of an overall navigation solution. Stalder et al. (2008) proposed a landmark detection scheme for future navigation work that is based on interferences in the local texture field. Aulinas et al. (2011) considered bright spots with dark shadows associated with potential reference point components to detect robust features from underwater environments using SUFR feature detection and a matching algorithm. On the navigation side, a number of papers have appeared that document attempts to use side-scan sonar imaging for SLAM (He et al. 2012; Pinto et al. 2009; Woock & Frey 2010). Similarly, other papers have described research that involved using bathymetric data to perform terrain-relative navigation (Claus & Bachmayer 2014; Meduna et al. 2010). Furthermore, Padial et al. (2013) showed that the acoustic shadows produced by an object projecting from the seabed can be used as a measurement of correlation between the observed shadows and path predicted from bathymetry for terrain-relative navigation. Furthermore, King et al. (2017) have shown that AUV navigation is a crucial problem in the appearance of the sea floor which is changing rapidly all the time

due to tides and currents. Moreover, long linear features along the vehicle deployment path caused another problem in the image-based AUV localisation.

Extracting good features of the environment is a highly complex problem, especially when it comes to underwater environments or under ice. This problem is quite pronounced when dealing with an under-ice environment with largely low contrast and lack of features (Spears et al. 2015). Thus, extracting key features in such an environment is the major challenge in AUV localisation. This becomes even more difficult in long-range AUV navigation. Therefore, AUV navigation and localisation in such terrain is virtually impossible without advanced feature detection methods. A number of feature detectors and descriptors have been used in the literature, such Harris Corner (Harris & Stephens 1988), Smallest Univalue Segment Assimilating Nucleus (SUSAN) (Smith 1992), Scale-Invariant Feature transform (SIFT) (Lowe 1999), Speeded-Up Robust Features (SURF) (Bay et al. 2006), Binary Robust Independent Elementary Features (BRIEF) (Calonder et al. 2010), Binary Robust Invariant Scalable Keypoints (BRISK) (Leutenegger et al. 2011) and Fast Retina Keypoints (FREAK) (Alahi et al. 2012). While these work well in complex environments, detectors and descriptors have reduced efficiency in featureless environments. Aiming to address this issue, advanced feature extracting algorithms such as combination of detectors and descriptors have to be used for under-ice environments.

2.4 Summary

In summary, a review of recent advances in under-ice localisation and navigation has been performed in terms of various sensors and algorithms along with their pros and cons. As discussed earlier, there is a common problem in data association response in SLAM and TRN. There has been a large amount of research conducted to realise the high accuracy sensors used in the under-ice environment. Moreover, the unavoidable growth of drift in position estimation is a major challenge in AUV navigation, especially in under-ice navigation. Therefore, mature technology side-scan sonars are the best solution for this issue as they use features in the environment which are naturally available for estimating the vehicle position. The key challenge in side-scan sonar images is insufficient high-low contrast and lack of significant features. Therefore, identifying reliable key features is a challenging task with current detectors used in feature detection. Aiming to address this problem, as a substitute for using a single detector, a combination of several detectors might detect more features in such environments. Therefore, this research makes contributions in the areas of sonar image generation, planned feature matching algorithms and proposed image based localisation system.

Chapter 3 Performance Comparison of Keypoint Detectors and Descriptors

3.1 Introduction

Localisation is a crucial issue given the dynamic and unstructured characteristics of underwater environments and sensors with high resolution and accuracy are needed in order to determine the correct track of the vehicle. In image based navigation, features in the environment are compared with stored maps with features. Optical imaging devices such as monocular or stereo cameras and sonars are the sensors that are most used in image based navigation. Visual udometry is mainly achieved through optical flow or structure of motion (SFM). Moreover, side-scan sonar images can be interpreted in a similar manner as optical images, although these images have very unique properties such as the presence of acoustic shadow (Nadir), resolution and range-varying attenuation (Blondel 2010). However, due to the limitations of optical cameras in underwater environments such as inadequacy of light, susceptibility to scattering, reducing range and blurred images due to the vehicle movements they are not popular in AUV navigation (Paull et al. 2014; Burguera et al. 2015).

As discussed in Chapter 2 of this thesis, the use of side-scan sonar images is a more feasible method for identifying the position of an AUV as it uses natural features available in the environment. Therefore, sonar image based AUV localisation is a promising solution as it requires neither deploying sensors nor the calculation of distance travelled from a reference point in order to determine location. In recent years there has been increased research interest in sonar imaging based AUV localisation in academia as well as in the industry (Bandara et al. 2016; Li et al. 2017; Song et al. 2016). Therefore, localisation and recognition of objects based on local point features has become a broadly used methodology.

However, the available feature detector and descriptor algorithms are originally designed for surface or air navigation. To date, very few groups have reported the application of these algorithms in underwater environments (King et al. 2013; King et al. 2014; King et al 2017; Nguyen et al. 2012). Aulinas et al. (2011) considered bright spots with dark shadows associated with potential reference point components to detect robust features from an underwater environment using SUFR feature detection and matching algorithm. Moreover,

Nguyen et al. (2012) highlighted the importance of using advanced spatial keypoint process and segmentation process for feature detection and matching in future development. Therefore, it is important to identify reliable feature detectors and descriptors which can work with available sonar images with different environmental conditions. As reported in the literature, SURF and Harris are better in repeatability (Fraundorfer & Scaramuzza 2012) while FAST is in fact the fastest in feature detection matching the best for efficiency (Rosten & Drummond 2006). Similarly, Harris is the best for localisation accuracy (Fraundorfer & Scaramuzza 2012). BRISK shows better immunity to noise compared to BRIEF (Schmidt & Kraft 2015). Furthermore, a preliminary study showed that SIFT is several times lesser in computation time compared to SURF and its computation complexity (Bay et al. 2006). Moreover, King et al. (2013) demonstrated, through experiment results, that SURF works comparatively better with SIFT with underwater sonar images. BRIEF is sensitive to image rotation and scaling (Schmidt et al. 2013). SURF, BRISK and FREAK are the best for robustness (Fraundorfer & Scaramuzza 2012; Schmidt & Kraft 2015). Therefore, this chapter presents a comprehensive performance comparison on keypoint detection and matching of SURF, Harris, FAST, BRISK and FREAK algorithms in underwater environments. However, rather than using optical imagery, the comparison will be based on side-scan sonar imagery (i.e., images of acoustic backscatter) of the seabed.

3.2 Detectors and Descriptors

Generally, many algorithms which come from computer vision depend on the determination of keypoints (interest points) in each image and calculating a feature description from the pixel region surrounding the keypoint. Terminology varies across the literature in feature detection and feature description. According to some discussions (Hassaballah et al. 2016; Krig 2014a), interest points are referred to as keypoints and the algorithms used to find the interest points are called detectors. Similarly, the algorithms used to describe the features are referred to as descriptors.

There are various types of detector methods available such as Laplacian of Gaussian (LOG), Harris and Stephens corner detection, Difference of Gaussian (DOG), Determinant of Hessian (DoH), Salient Regions, SUSAN, and Morphological interest points. Each detector has its own advantages and disadvantages when working in different environments. In the same way, there are various descriptors identified in the literature such as Local Binary descriptors, Spectra descriptors and Basic Space descriptors and Polygon Shape descriptors (Krig 2014a). Furthermore, SURF, Harris, FAST, BRISK and FREAK use their own detector and descriptor which will be discussed in detail in the next section; as an example, SURF uses Determinant of Hessian (DOH) as its detector and its descriptor is based on 'Harr wavelet' which is based on the Spectra descriptor (Bay et al. 2006). The success of sonar image processing based localisation heavily depends on the performance of its detector and descriptor. The feature detection and matching algorithms that are used in this study are briefly explained below.

3.2.1 Speeded-Up Robust Features (SURF)

The SURF algorithm consists of a feature detector and a descriptor algorithm, which is inspired by the SIFT algorithm (Schmidt et al. 2013). It computes the Hessian matrix given in Eq. (3.1) as the output of the detector. The original authors of this method claim that the Hessian matrix yields better performance in terms of computation time and thus achieves fast detection of features (Bay et al. 2006). The Hessian matrix $H(\mathbf{x}, \sigma)$ of image *I,* given point of $\mathbf{x}$ = *(x, y)* and scale σ is defined in Eq. (3.1):

$$H(\mathbf{x}, \sigma) = \begin{bmatrix} L_{xx}(\mathbf{x}, \sigma) & L_{xy}(\mathbf{x}, \sigma) \\ L_{xy}(\mathbf{x}, \sigma) & L_{yy}(\mathbf{x}, \sigma) \end{bmatrix} \tag{3.1}$$

where $L_{xx}(\mathbf{x}, \sigma)$, $L_{xy}(\mathbf{x}, \sigma)$ and $L_{yy}(\mathbf{x}, \sigma)$ are the convolution of the Gaussian second order derivative with image *I* in point *x,* point *xy* and point *y* respectively. The approximated Hessian determinant symbolises the blob response of the image point at $\mathbf{x}$. A detailed description on SURF can be found at Bay et al. (2006).

The descriptor part of the SURF algorithm produced an intensity contained distribution based on the *x* and *y* directions of 'Harr wavelet' responses within the neighbourhood area of an interest point (Bay et al. 2006). 'Harr' features are based on specific sets of rectangular patterns which approximate the basic 'Haar wavelets' and each 'Haar' feature is able to be processed using the average pixel value of the pixels within the rectangle (Krig 2014a). Owing to the simplicity and invariant properties against scale and rotation, SURF is considered to be a fast and robust feature detection algorithm (King et al. 2013).

3.2.2 Harris Detector

The Harris detector algorithm calculates a matrix in relation to the autocorrelation function of the image. The corresponding mathematical representation is given in Eq. (3.2) which results in two eigenvalues of the autocorrelation function from its major curvatures (Harris & Stephens 1988). If one eigenvalue is positive, that indicates an edge while a corner results in positive values for both eigenvalues. Harris points are invariant to the rotation but vary with scaling.

According to the authors (Harris & Stephens 1988), the sum of the squared difference (SSD) is used for calculating the autocorrelation function for corner detection $E(x, y)$ of an image. In image I which has an image patch (u, v) over the area and shifted it by (x, y), which represents the difference in intensity for a displacement in all directions. Similarly, $w\ (u, v)$ represents the window function and $I(x + u, y + v)$, $I(u, v)]^2$ represents the shifted intensities and intensities of the function respectively. A detailed description of the Harris detector can be found in Harris & Stephens (1988).

$$\boldsymbol{E}\left(\boldsymbol{x}, \boldsymbol{y}\right) = \sum_{\boldsymbol{u},\boldsymbol{v}} \boldsymbol{w}\left(\boldsymbol{u},\boldsymbol{v}\right)[\boldsymbol{I}(\boldsymbol{x}+\boldsymbol{u}, \boldsymbol{y}+\boldsymbol{v}) - \boldsymbol{I}\left(\boldsymbol{u},\boldsymbol{v}\right)]^2 \tag{3.2}$$

3.2.3 Features from Accelerated Segment Test (FAST)

The Features from Accelerated Segment Test (FAST) algorithm is mainly a corner detector which is capable of extracting feature points as well. This algorithm is focused on determining a 'corner' and is dependent on 16 pixels which form a circle around the candidate point. If there are 12 contiguous pixels, all of which are brighter than or darker than the circle, the candidate point along with the threshold value is accepted as a keypoint. This is carried out by the segment test which is considers a circle of 16 pixels around the corner candidate p. The original classifies p as a corner sensor, if there is a set of n contiguous pixels in the circle, which are brighter than the intensity of candidate pixel I_p plus a threshold t, or all darker than $I_p + t$, as illustrated in Figure 3-1.

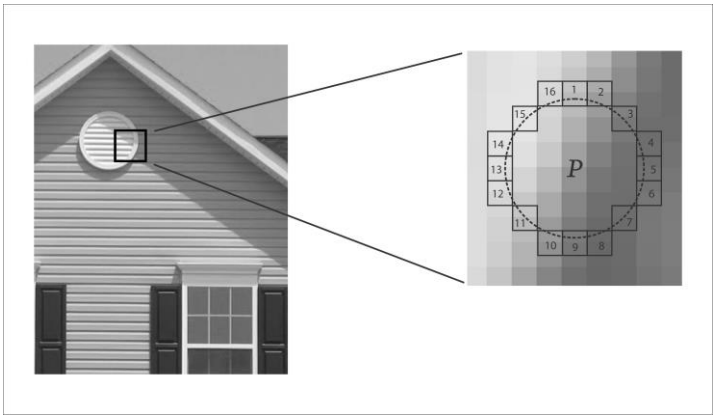


Figure 3-1. The segment test used by the FAST descriptor (Rosten & Drummond 2006)

As the name suggests (although perhaps coincidental), this algorithm is significantly faster in comparison to all other feature extraction methods and is therefore becoming popular in robots and AUVs that have limited computing resources (Rosten & Drummond 2006). A detailed description of FAST is found in Rosten & Drummond (2006).

3.2.4 Binary Robust Invariant Scalable Keypoints (BRISK)

The Binary Robust Invariant Scalable Keypoints (BRISK) algorithm is a fast feature detection, description and matching algorithm. The main reason behind its speed is a FAST-based scale-space novel detector. Keypoints are detected in octave layers of the scale-space pyramid as shown in Figure 3-2. The octaves are formed by progressively half sampling the original image. Each layer of the pyramid consists of n layers, octaves of Ci and intra-octaves of di for $i = \{0, 1, \ldots\ldots\ldots n-1\}$, where n is normally considered as 4. This detector is associated with combining a bit-string descriptor for intensity comparing for its sampling pattern in the nearest keypoints (Leutenegger et al. 2011). As per its name, BRISK is capable of rotation as well as scale invariance to an important extent.

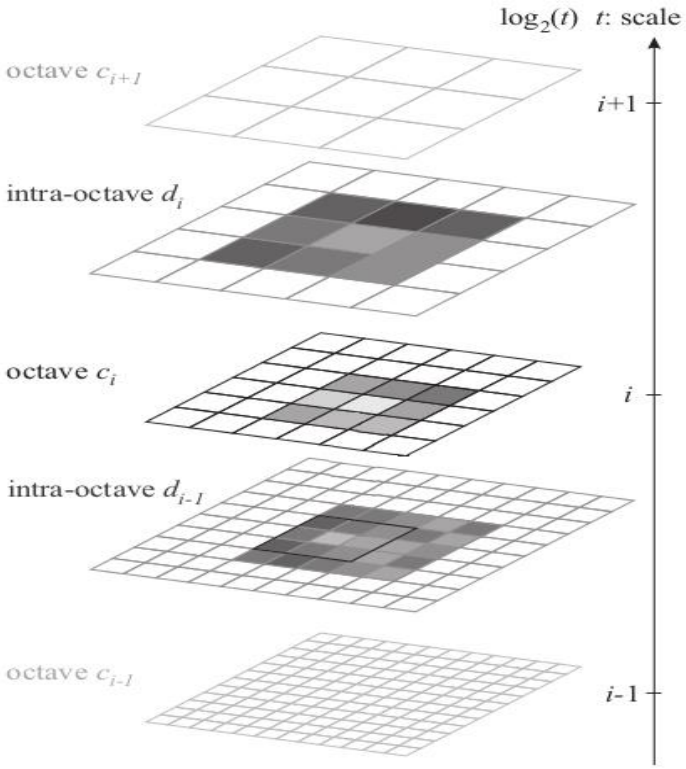


Figure 3-2. Scale-space keypoints detection of BRISK (Leutenegger et al. 2011) © [2011] IEEE

The BRISK descriptor uses a sampling pattern of neighbourhood of keypoints as illustrated in Figure 3-3. A detailed description of BRISK is given in Leutenegger et al. (2011).

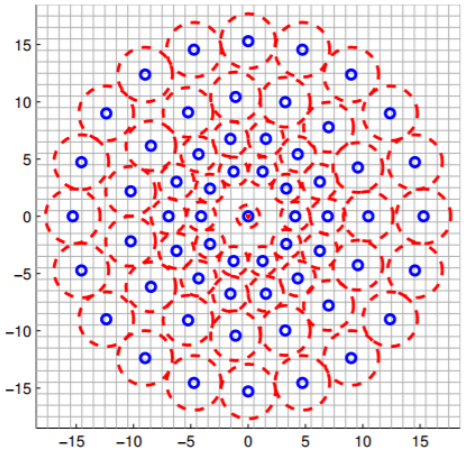


Figure 3-3. Sampling Pattern of BRISK descriptor (Leutenegger et al. 2011) © [2011] IEEE

3.2.5 Fast Retina Keypoints (FREAK)

The Fast Retina Keypoints (FREAK) algorithm is a binary feature descriptor algorithm inspired by the visual system of humans, which is indeed the retina. Retinal sampling pattern compares image intensities by a cascade of binary string and higher density of sampling points closer to the centre. After sampling, the binary FREAK descriptor is constructed from a one-bit series of Difference of Gaussians (DoG). The Saccadic search algorithm is performed subsequently to identify significant key features (Alahi et al. 2012). Finally, rotation keypoints available from each identified key feature are calculated using the calculation of local gradients over selected pairs. A detailed description of FREAK is available in Alahi et al. (2012).

As mentioned in the introduction to this chapter, this research carried out a comprehensive performance comparison of the above-mentioned five detector and descriptor algorithms with filtering effect, covering a wide range of scenarios that can occur in an underwater environment. This section represents the basic working principal of each detector and descriptor and evaluation of detector–descriptor pairs in the context of robot navigation.

3.3 Methodology

In this study, MATLAB 8.5.0 software was used to process the images. Relevant tools that are available in the MATLAB computer vision system library were used to prepare images, and detect and match keypoints. As shown in Figure 3-4, the image processing has four major steps, namely: (1) image acquisition, (2) image preparation, (3) keypoint detection and (4) keypoint matching. These steps are discussed in detail in the following sub-sections.

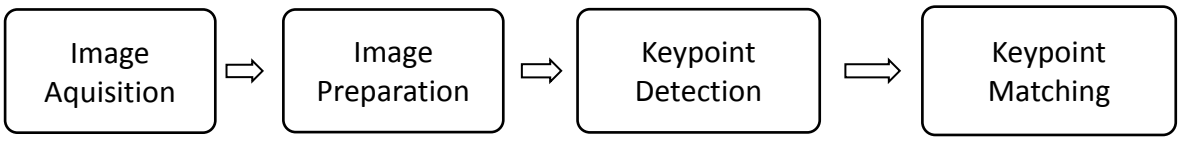


Figure 3-4. Major steps in image processing

3.3.1 Image Acquisition

The AUV Explorer deployed at Holyrood Arm, Newfoundland and Labrador, Canada in 2014 was used to acquire images. The AUV was operated in a constant altitude mode of 7 m from the bottom and maintained a constant speed of 1.5 m/s. A map of the surveyed area is shown in Figure 3-5, which is relatively flat containing sand, mud with some exposed rocks (<1 m in height) and source marks from anchors. The raw data were obtained from the AUV's integrated EdgeTech 2200M side-scan sonar in the native 'JSF' format. The sonar was mounted in the forward hull along the longitudinal axis of the AUV and operating with the data collected at 400 kHz.

The collected sonar data consist of a series of records, called 'pings'. There are 1001 pings in each tile and each tile represents 44 m of travel path along a 1 km line (King et al. 2012). The obtained raw data were processed as shown in Figure 3-6 to generate image files. These data were converted into a single coordinate system with respect to the vehicle's viewpoint (vehicle heading and speed) which is placed in a global coordinate system relative to true north. These adjustments mean that the path and reference images are always oriented north up (King et al. 2012). A complete description of raw image generation is found in previously conducted work (King et al. 2012). The vehicle's pitch and roll are assumed to

be constant as the Explorer AUV has six large control planes and can change its depth with negligible variation in its pitch and roll while undergoing straight line motions. Previous work conducted using the AUV Explorer has also shown that vehicle's pitch and roll to have negligible impact on the data quality (King et al. 2012).

Figure 3-5. Survey terrain in Holyrood Arm, Newfoundland and Labrador, Canada (47.388N, 53.1335W) (King et al. 2014)

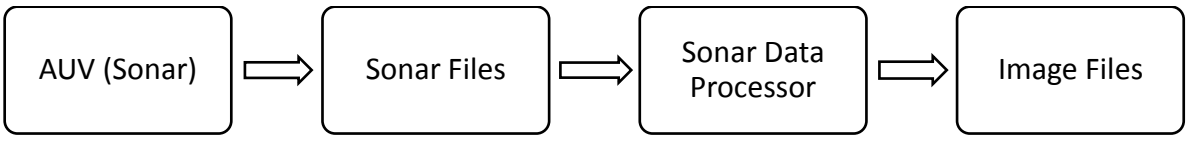
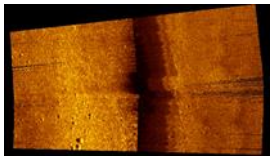
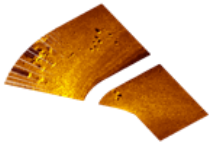
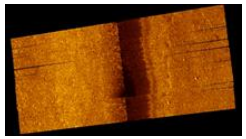
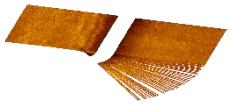


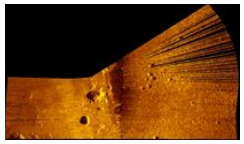
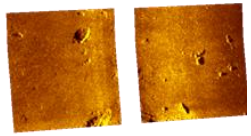
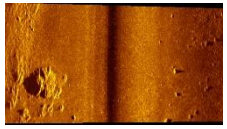
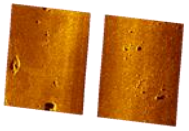
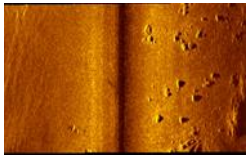
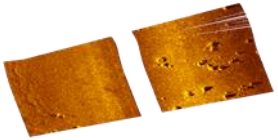
Figure 3-6. Data flow of sonar image capture

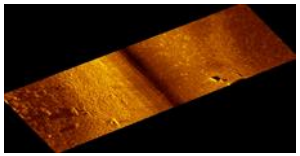
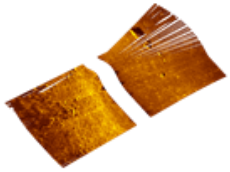
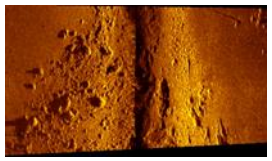
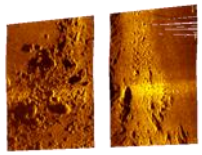
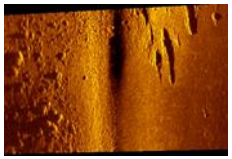
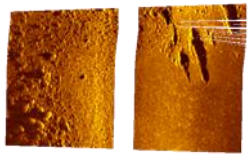
An additional survey was carried out in the same area at different time periods on the following day in order to obtain sonar images with variations compared to the images taken on the first day. The surveyed area consists of a wide range of bottom terrain scenarios such as clutter, mud, sand, stones and lack of features; and sonar imaging effects such as scaling, rotation, non-uniform intensity (difference in brightness) and backscatter. The difference in brightness was caused by the sonar beam pattern along with the ever-changing attitude of the tow-fish which create difference seabed images. During the deployment of the AUV, the side-scan sonar simultaneously acquired seabed images, and the acoustic received level in addition to imagery data from the sonar were able to be converted to estimate the backscattering strength of the seabed. Furthermore, a large-range side-scan sonar system is able to generate 'acoustic brightness' images of seabed swaths using the acoustic energy returned from the seabed (Chang et al. 2010). The 20 images obtained in this survey are grouped into 10 sets, representing 10 selected locations of survey a day apart and covering a variety of terrain scenarios and image effects as shown in Table 3-1. For example, Images 7 and 8 in Data Set 4 have rotation, scale, difference in brightness, fewer stones on the surface and images are not in the same area. Similarly, Images 9 and 10 in Data Set 5 have rotation, scale and more stones on the surface compared to Data Set 4. Amidst having these variations, the majority of the features found in the images obtained on Day 1 should match with the corresponding features in the images taken on Day 2 as they are taken from the same course of the AUV.

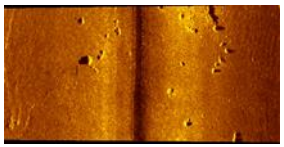
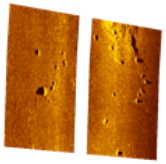
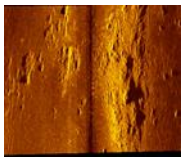
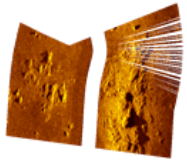
The black centre of the image sets from Day 1 represents the 'nadir' (i.e. the area directly below the vehicle which is not covered by the sonar). This area does not contain data and has no effect on the keypoint detection and keypoint matching process.

Table 3-1. Image data sets

Image Set	Images from Day 1	Images from Day 2	Effects/Features
Data Set 1	 (Image 1)	 (Image 2)	Rotation Scale Differences in brightness Muddy surface Less overlapping area
Data Set 2	 (Image 3)	 (Image 4)	Rotation Scale Differences in brightness Muddy surface Less overlapping area Lack of features

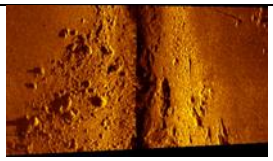
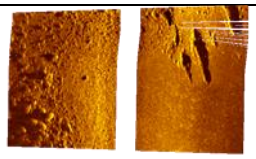
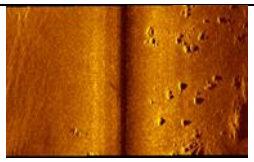
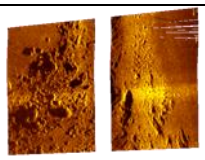
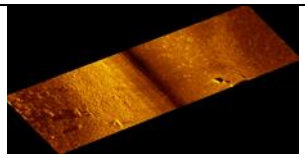
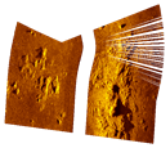
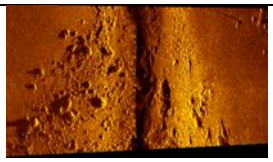
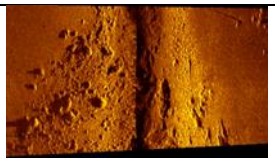
Data Set 3	 (Image 5)	 (Image 6)	Rotation Scale Differences in brightness Muddy surface with stones or sand Less overlapping area
Data Set 4	 (Image 7)	 (Image 8)	Rotation Scale Differences in brightness Muddy surface with stones or sand Not in the same area
Data Set 5	 (Image 9)	 (Image 10)	Rotation Scale Muddy surface with stones or sand More overlapping area

Data Set 6	 (Image 11)	 (Image 12)	Rotation Scale Differences in brightness Muddy surface with stones or sand More overlapping area
Data Set 7	 (Image 13)	 (Image 14)	Scale Tilt Clutter Muddy surface with stones or sand More overlapping area
Data Set 8	 (Image 15)	 (Image 16)	Scale Muddy surface with stones or sand More overlapping area

Data Set 9	 (Image 17)	 (Image 18)	Scale Mud surface with stones or sand More overlapping area
Data Set 10	 (Image 19)	 (Image 20)	Scale Distortions Clutter Muddy surface with stones or sand More overlapping area

In addition to the image sets shown in Table 3-1, another four image sets are created by combining some of the images from Table 3-1 to analyse the effects of scale and changes in the difference of brightness conditions in the environment which is not available in Table 3-1. The corresponding image sets are given in Table 3-2.

Table 3-2. Creation of data sets

	Images from Day 1	Images from Day 2	Effects/Features
Data Set 11	 (Image 21)	 (Image 22)	Scale Differences in brightness Muddy surface with stones or sand Less overlapping area
Data Set 12	 (Image 23)	 (Image 24)	Slightly different in scale Muddy surface with stones or sand Less overlapping area
Data Set 13	 (Image 25)	 (Image 26)	Rotation Scale Differences in brightness Muddy surface with stones or sand Less overlapping area
Data Set 14	 (Image 27)	 (Image 28)	Same Image

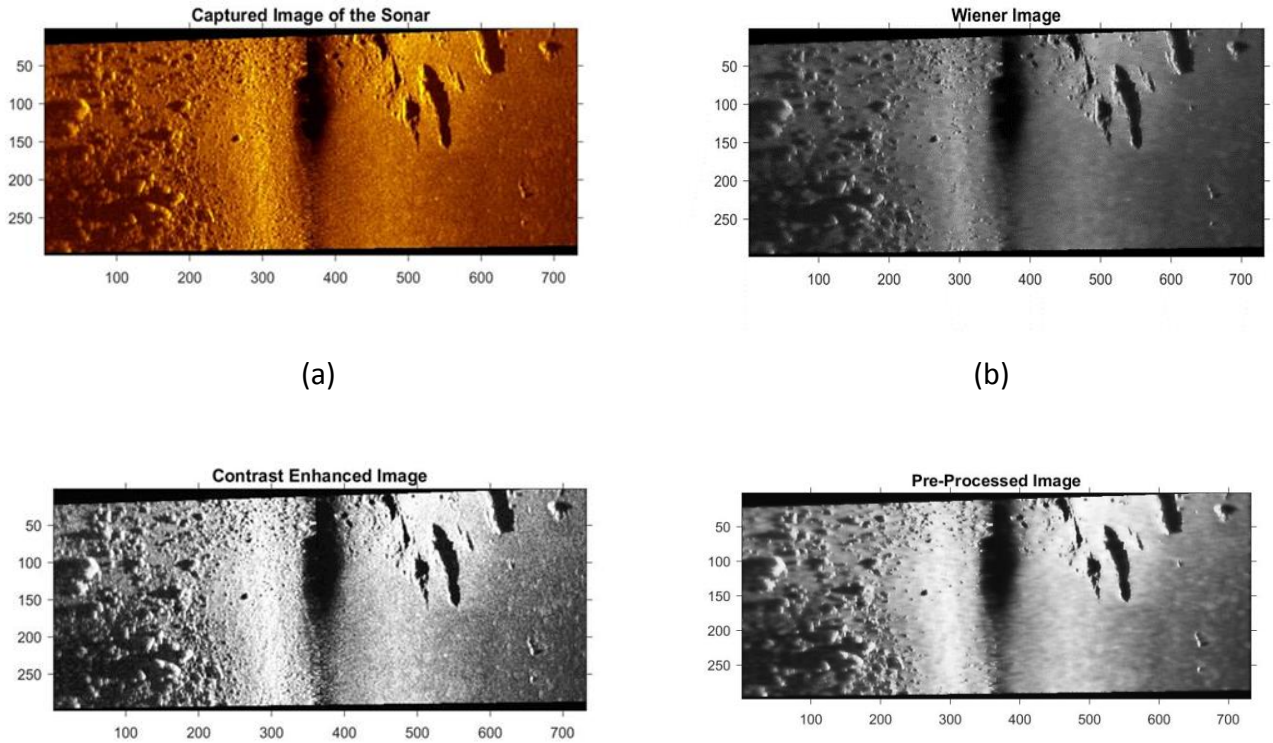
3.3.2 Image Preparation

The acquired images have different features and characteristics such as sand surface, mud surface with sand, mud surface with sand and stones, noise and blur. Therefore, to create uniformity across all the images, they need to be passed through a preparation stage which is divided into four sub-stages, namely: 1) resizing, 2) filtering, 3) contrast enhancement and 4) morphologic filtering. The effects of these stages are illustrated in Figure 3-2 taking Image 15 of Data Set 8 as an example. As shown in Figure 3-7(a), the sonar images acquired by the sensors are 732 x 299 pixels in size. When the acquired image is not the above-mentioned size, the image is resized to 732 x 299 pixel size.

These images are then passed through a Wiener filter to remove any noise generated by electronic or mechanical systems inside the sonar as shown in Figure 3-7(b) (Wiener 1949). Generally, median filters, average filters, Gaussian filters and Wiener filters are used for image enhancement in image processing. Furthermore, non-linear median filters are commonly known for removing 'salt and paper' noise. In the same way, average filters or median filters are simple filters that are good for noise removal although they present problems of blurring in the image (Arya and Semwal, 2017). Gaussian filters are known for blurring and suppressing the image. Therefore, the main reason for selecting the Wiener filter for this process is due to its inherent capability of minimising the mean square error in the process of smoothing the noise and inverse filtering. Furthermore, according to the literature, the Wiener filter is capable of holding a high Peak Signal to Noise Ratio (PSNR) value which indicates that the image holds high qualify even after de-noising (Williams et al. 2016).

The next step is contrast enhancement which develops the perceptibility of the image in such a way that the result of brightness difference between objects and the background is much better than the original image for a particular application (Gonzalez & Woods 2002). At this stage, the image contrast was enhanced by computing a histogram of each level of intensity in a given image and stretching it to have a thinner range of intensities, as illustrated in Figure 3-7(c). This particular manipulation produces an image with higher contrast. After the enhancement, as shown in Figure 3-7(d), a series of morphological filters are used to filter out any noise introduced by the enhancement process (Gonzalez & Woods 2002; Marques 2011). Furthermore, when the image becomes blurred in comparison to the original image,

this helps the keypoint extraction to focus on larger regional keypoints. Therefore, there is no need to add images to fill gaps in the samples as the blurring naturally fills these gaps (King et al. 2012).



(a) (b)

(c) (d)

Figure 3-7. (a) Original image before resizing and filtering; (b) image after filtering; (c) contrast enhanced image and (d) morphologically filtered image with Data Set 8 first day image (Image 15)

3.3.3 Keypoint Detection

As depicted in Figure 3-3, keypoint detection is the third step of the image processing workflow. At this stage, the pre-processed image is passed through the detection algorithm in order to extract keypoints. The algorithm identifies regions of the image that have unique content, such as corners or blobs (image pattern of region which are different from their surrounding intensities), depending upon the algorithm used. The important factor in the determination of the keypoint is to find key features that remain locally invariant so that it is possible to detect these key features even in the presence of rotation or scale change. Depending on the image, the algorithm used and threshold values set in the algorithm, the number of keypoints detected can be as high as 1000 in feature-rich images, which may load the processor heavily in the next stage.

On the other hand, images that lack features may result in zero or a very low number of keypoints. In both cases, the variable used for a given algorithm is the detection threshold. For feature-rich images the threshold can be raised to extract only the prominent keypoints. The obvious disadvantage is the loss of important points that may work well in subsequent stages. Similarly, for lack of features with low-contrast images, the threshold can be lowered to extract more keypoints. The drawback of lowering the threshold is the possibility of false detection. Therefore, choosing the right threshold is important in order to achieve fast and robust matching of features in sonar images (Gonzalez & Woods 2002).

In order to achieve uniformity across all the methods compared, 1000 was set as the detection threshold in this study for all the data sets (14 data sets). This particular value was chosen by trial and error. Generally, thresholding methods can be divided in to global and local thresholding and each method can be further subdivided as default, Huang, Otsu, mean, median etc. (Krig 2014a). Furthermore, one of the key challenges to be faced in calculating the threshold is the non-uniform illumination of the images which was handled with image pre-processing. Once the keypoints are detected, corresponding descriptors are calculated by the algorithm used. The detected keypoints and their descriptions are then stored in an m x n matrix form and passed to the next stage for keypoint matching (Matlab 2017).

3.3.4 Keypoint Matching

The last step of sonar image processing is the feature matching where the detected features are compared against the features in the images stored in a matrix form described in section 3.3.3. In practical implementations, this comparison should be carried out against all the images recorded for the particular area of interest. Nevertheless, as mentioned in section 3.3.1, 14 image sets, each having two images, taken a day apart and of the same area, are chosen for comparison in this study. Therefore, features of the two images in a set should match with each other.

In order to match keypoints detected in one image with those in the other image there should be a unique description for each point. The algorithms developed that generate such descriptions are called 'descriptors'. The feature descriptors are mainly categorised as Local Binary descriptors, Spectra descriptors and basic space descriptors (Krig, 2014b). In this study, SURF, BRISK and FREAK descriptors were used. The feature descriptor of SURF is based on the sum of the 'Haar' wavelet response within the neighbourhood area of a keypoint (Bay et al., 2006). This descriptor basically belongs to the category of Spectra descriptor that is capable of computing light intensities, colour, local area gradients and local area static features. On the other hand, BRISK and FREAK belong to the category of local descriptors that are efficient and competent in computing, storing, and matching features (Krig, 2014b). The BRISK descriptor compares the binary outcome pattern of keypoints as discussed in Section 3.2.4 (Leutenegger et al., 2011). Furthermore, FREAK is a binary descriptor that operates on a cascade of binary strings while calculating the image intensities that end on over retina sampling pattern (Alahi et al., 2012). However, Harris and FAST are feature detectors and not performing as a feature descriptor. Nevertheless, FAST determines the 'corner' after comparing all the pixels in a circular pattern against the central pixel using a threshold value. Therefore, the resulting descriptor is stored as a contiguous bit vector in order of 1-16 in a rotational-invariant representation pattern (Krig, 2014b).

In the matching process, those descriptions are compared and the points with minimum deviations are taken as matches. The deviations are measured as the Euclidean distance (vector magnitude). The commonly used matching strategy of the Euclidean distance is to set a threshold value and return all matches of the other image within this threshold

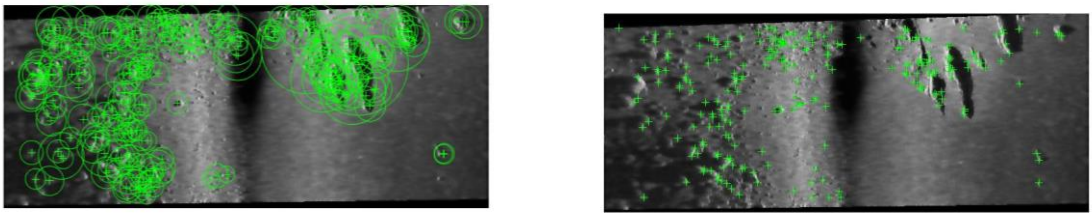
value. This threshold value can be adopted in different regions of feature space. Therefore, to maintain the unique threshold value for all image data sets, the distance match of the nearest neighbour is used.

Therefore, this matching process calculates the nearest neighbours between the identified features of the two images. If the best match's distance is less than a definite match of threshold value, the match is retained. Otherwise, the match is rejected. In this study, the distinct match threshold value was set to 10.0 for binary feature vectors and 1.0 for non-binary feature vectors. Moreover, EstimateGeometricTransform has been used to get the best match in matching features. EstimateGeometricTransform is matching comparatively the set of consecutive images. M-estimator Sample Consensus (MSAC) algorithm is used to remove outliers in corresponding matched features which are a variant of the Random Sample Consequence (RANSAC) algorithm.

3.4 Results and Discussion

3.4.1 Keypoint Detection

As discussed in Section 3.3, keypoints have been detected from all the pre-processed Data Set images using SURF, Harris, FAST and BRISK detectors. Figure 3-8 shows the example of the above-mentioned keypoint detection for SURF and FAST algorithms on Image 15 of Data Set 8 (first-day image).



(a) (b)

Figure 3-8. Keypoints detected in Data Set 8 (Image 15) with the (a) SURF algorithm; (b) FAST algorithm

Similar to the Image 15 shown in Figure 3-6, each image and its filtered version were passed through all the algorithms used in this study and the number of keypoints detected was tabulated as in Table 3-3. An interesting observation that can be made through the results listed in Table 3-3 is that the original images contain more keypoints than the filtered images. Therefore, with the aim of analysing the effect of filtering, the retained percentage of the keypoints is also examined (see Table 3-3). The retained percentage indicates what percentage of features detected in the original image is retained when it is filtered. The corresponding plot is shown in Figure 3-9 which leads to an interesting observation where SURF is able to maintain the retained percentage within a stable region compared to large variations present in other algorithms. This stable region is a user-defined region of retain percentage from 40% to 60% derived from the keypoint matches from Table 3-3. This shows that although the keypoint detection capability of SURF is lower than Harris or FAST, SURF is able to maintain stable keypoint detection for all the data sets compared to other detectors, due to robust detector capability.

Table 3-3. Number of keypoints detected by each algorithm

Image Data Set	Image No.	Day	SURF			Harris			FAST			BRISK		
			Without filter	With filter	Retain (%)	Without filter	With filter	Retain (%)	Without filter	With filter	Retain (%)	Without filter	With filter	Retain (%)
1	1	1	252	107	42	731	280	38	592	228	39	278	121	44
	2	2	278	222	80	500	394	79	687	522	76	323	264	82
2	3	1	255	135	53	541	362	67	398	222	56	165	105	64
	4	2	364	282	77	872	794	91	900	719	80	265	206	78
3	5	1	472	260	55	1061	657	62	938	430	46	326	192	59
	6	2	203	147	72	257	186	72	258	162	63	151	121	80
4	7	1	222	167	75	561	317	57	247	107	43	143	99	69
	8	2	169	106	63	302	224	74	284	191	67	161	132	82
5	9	1	148	101	68	233	141	61	212	122	58	147	114	78
	10	2	214	148	69	400	343	86	327	235	72	160	151	94
6	11	1	193	124	64	264	178	67	367	211	57	164	101	62
	12	2	377	282	75	701	502	72	959	607	63	361	263	73
7	13	1	649	393	61	1002	441	44	995	342	34	558	275	49
	14	2	600	411	69	1021	567	56	1031	463	45	558	328	59
8	15	1	438	259	59	545	225	41	658	248	38	330	199	60
	16	2	466	267	57	736	342	46	833	315	38	424	223	53
9	17	1	110	66	60	127	90	71	171	103	60	115	94	82
	18	2	204	150	74	306	220	72	249	166	67	157	131	83
10	19	1	460	315	68	665	323	49	623	232	37	331	188	57
	20	2	525	385	73	905	587	65	952	545	57	455	339	75

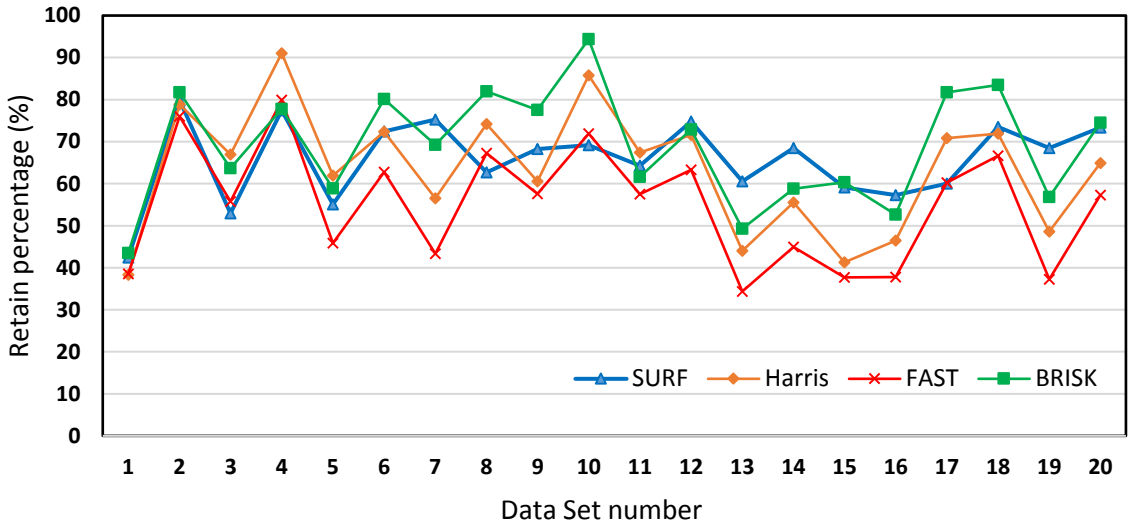


Figure 3-9. Retain percentage of the number of key points

As shown in Table 3-4, the Harris algorithm detects more keypoints followed by the FAST algorithm, which failed to perform in keypoint matching after analysing both tables. The results in Table 3-3 further indicate that SURF is more on the lower side in keypoint detection while BRISK produces the lowest number of keypoints in most of the cases. Although SURF detects fewer keypoints than Harris and FAST, it is able to maintain the highest number of matches. Furthermore, SURF is able to perform better than FAST in Data Sets 7, 8 and 10 due to the cluttering of feature patterns in images.

When two images are taken for matching, the maximum possible number of matches is decided by the lower number of keypoints detected in one of the two images. For example, the SURF algorithm detects 259 and 267 keypoints in Image 15 and Image 16 respectively. The maximum possible number of matches between the two images is 259. Similarly, for each image set, the lower number is taken as the maximum possible number of matches and the corresponding listing is shown in Table 3-4. The same table shows the number of matches and the matching percentage for each set and each algorithm. For example, although the image Data Set 8 has the possibility of having 259 matches, the actual number of matches is 37 which is equal to 14.3% matching. The corresponding matches are shown in Figure 3-10.

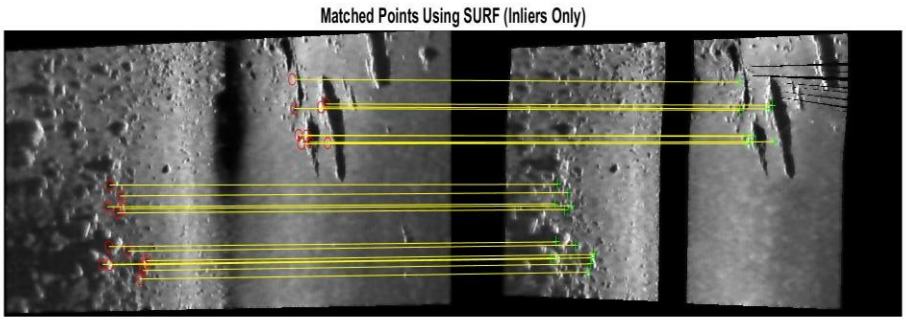


Figure 3-10. Key points matching using SURF Data Set 8

Table 3-4. Maximum possible number of keypoint matches and actual number of matches for each algorithm

Image Data Set	SURF			Harris			FAST			BRISK			FREAK		
	Keypoints	Matching	Percentage (%)	Keypoints	Matching	Percentage (%)	Keypoints	Matching	Percentage (%)	Keypoints	Matching	Percentage (%)	Keypoints	Matching	Percentage (%)
1	107	0	0	280	0	0	228	0	0	121	0	0	121	0	0
2	135	0	0	362	0	0	222	0	0	105	0	0	105	0	0
3	147	0	0	186	0	0	162	0	0	121	0	0	121	0	0
4	106	0	0	224	0	0	107	0	0	99	0	0	99	0	0
5	101	5	4.95	141	1	0.71	122	3	2.46	114	0	0	114	2	1.75
6	124	2	1.61	178	0	0	211	1	0.47	101	0	0	101	0	0
7	393	28	7.12	441	1	0.23	342	7	2.05	275	4	1.45	275	3	1.09
8	259	37	14.3	225	6	2.67	248	7	2.82	199	8	4.02	199	5	2.51
9	66	16	24.2	90	3	3.33	103	3	2.91	94	7	7.45	94	3	3.19
10	315	30	9.52	323	8	2.48	232	6	2.59	188	8	4.26	188	6	3.19

The number of actual matches and the mapping percentage achieved by each algorithm for different image Data Sets are plotted in Figures 3-11 and 3-12 respectively. As visible to the naked eye, image Data Sets 1–4 and 11–13 do not have matching features, which are confirmed by the results shown in Figure 3-11. Similarly, amidst rotations and scaling, the image Data Sets 5–10 can be seen as matching sets for naked eye with clearly distinguishable features. This is confirmed by the results having a varied number of matches produced by each method. The results in Figure 3-11 further reveal that the SURF algorithm performs well in all the cases, producing the highest number of matches. Figure 3-12 further consolidates the performance of SURF over other methods by plotting the matching percentage.

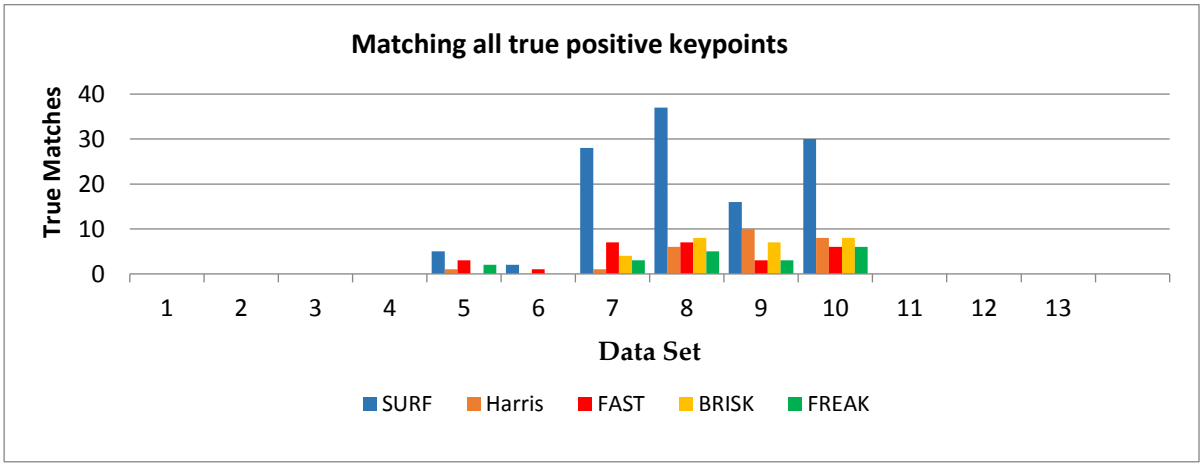


Figure 3-11. Number of matches produced by each algorithm for Data Sets 1–13 (note: Data Set 14 is deliberately removed from this diagram as it matches the features in the same image)

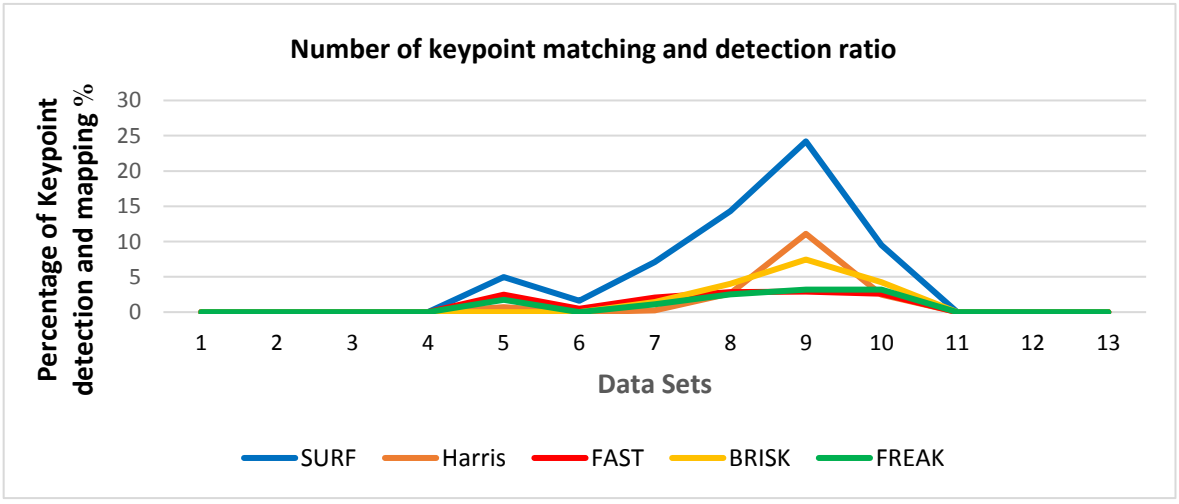


Figure 3-12. Number of keypoints matching percentage

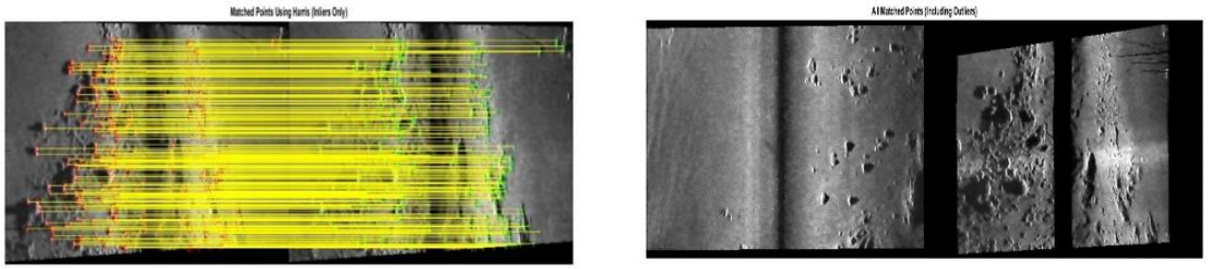
After comparing both Data Sets 5 and 6 in Figures 3-11 and 3-12, Data Set 5 is able to keep a higher number of keypoints matching for the SURF descriptor due to rotation and scale without distortion image. These two data sets have fewer sparse arrangement features. Moreover, the FAST detector performs in similar way to SURF. However, Harris and FREAK are unable to perform a single matching in Data Set 6 due to brightness as these algorithms fail to work in the discipline of brightness. Furthermore, BRISK is sensitive to image rotation and fails to work in all rotated images.

Further analysis of Figures 3-11 and 3-12, corresponding to the results of Data Sets 7–10, reveals not only the presence of features but also the distribution pattern of features affect the matching process. As can be seen by the naked eye, the images of Data Sets 7, 8 and 10 contain more features compared to Data Set 9. Nevertheless, they are cluttered features which can be described as a number of features appearing in a sonar image, where these features are quantified in terms of locally similar features covered untidily at some places in the images as illustrated in Data Sets 7, 8 and 10 in Table 3-1. In contrast, Data Set 9 contains sparse features which are spread throughout the entire image. The effect of clutter and distribution of features is clearly visible in Figure 3-12 where cluttered images produce a relatively lower percentage of matching compared to the high percentage of matching with distributed features. This observation is valid for all the algorithms compared. Nevertheless, SURF performs well in all the cases even with rotation, scale and differences in brightness. FAST and FREAK work well against rotation compared to Harris and BRISK. Harris is the least effective performer for images with rotation.

Moreover, the number of keypoints detected and the number of keypoints matched are very low, as given in Table 3-4. The highest number of keypoints mapping ratio was found in SURF as 24.24% for image Data Set 9. This shows that the important factor of increasing the number of the ratio of keypoint detector to keypoint matching should be done in future developments. This leads to an important observation where SURF produces the maximum percentage of matching even if the number of keypoints is low. Therefore, based on these results, two observations can be made: 1) although SURF does not detect more keypoints, its detection is robust, and 2) even if there are fewer keypoints SURF produces a higher percentage of matching.

3.4.2 Verification of Keypoint Matching

There is a major identification in Data Set 14 in Table 3-2, which gives 100% matching for all the keypoint detectors and descriptors as it compares the same image with its own features as shown in Figure 3-13(a). Similarly, Data Set 12 does not produce any matching as the two images are from two different places. This shows that the matching algorithms work as expected. During the matching process, images are generated without a single matching, and these are called non-matching images as shown in Figure 3-13(b).



(a) (b)

Figure 3-13. (a) Harris detector with all of the keypoints matching; (b) None of the keypoint matching

Moreover, another key observation that can be made through these results is that low contrast and environment with lack of features, as shown in Data Sets 1 and 2, are not able to perform well with SURF, Harris, FAST, BRISK and FREAK algorithms. If the entire surveyed region lacks image features, it is difficult to identify the key features with these feature detectors and descriptors. These keypoint detectors and descriptors are relatively well studied in land terrain as compared to underwater. However, the major challenge of underwater sonar images comes from the unstructured nature of the terrain with mainly low contrast and having lack of features or being featureless. Therefore, extracting proper key features in particular environments is the key challenge without having an advanced feature detection method.

The key challenge in image-based navigation is identifying robust key features in underwater terrain which could be low contrast, featureless and unstructured. Therefore, extracting good features in such environments is not possible with current feature detecting

methods such as SURF, Harris, FAST, BRISK and FREAK. According to the results, it is proved that these detectors and descriptors have to be more robust to work in such environments. Therefore, image-based localisation in such terrain is almost impossible without having an advanced feature detection method. Moreover, instead of using a single keypoint detector or descriptor, it is better to use their possible combinations for enhanced results. This can be done by using the most significant factor in each detector or descriptor. For example, if the Harris detector is the best for keypoint matching, use this particular detector as the detector component 1 and use another detector as detector component 2 because it is the second highest for finding keypoints. After that, the keypoints of the two components can be calculated as feature detector and find out the descriptor that is capable of handling robust environment. This combination of component 1 and 2 detector and the descriptor can be used to find keypoints in underwater or under-ice environments that lack significant texture and distinguishing patterns.

3.5 Summary

This chapter has presented the performance comparison of SURF, Harris, FAST, BRISK and FREAK algorithms adding filtering effect in keypoint detection and matching of sonar images. The reported work shows that each detector and descriptor is capable of handling images from various disciplines such as rotation, scale and brightness. In addition to that these detectors and descriptors work with sonar images in four different feature categories such as a low number of features with clutter, a low number of features with sparse arrangement, a high number of features with clutter and a high number of features with sparse arrangement. These results prove that SURF performs well in all the data sets where BRISK is failing to work images with noise and rotation. However, FAST managed to work with all the data sets after SURF. Harris is suitable for feature detection as compared to the mapping process. FREAK performed well in rotation and scale images and failed to work with brightness environment of images due to changes in the image intensity for retina sampling pattern.

The major conclusion that can be derived from the results is that the SURF algorithm outperforms all the other methods in keypoint matching with filtering even in the presence of scaling, rotation differences in image intensity. SURF also works well in all the disciplines with higher percentage matching even though it produces fewer keypoints, thus demonstrating its robustness. Furthermore, it has been shown that the number of detected keypoints reduced while using filtering effect. Nevertheless, it has been proved that the number of true key matches increased with filtering. It is also concluded that the distribution of features within the image has a significant impact on the matching process. Even if there is a large number of features in a cluttered environment, it produces less matching compared to features that are distributed. Therefore, another conclusion that can be derived from the results is that the feature detection and matching algorithms performed well in discrete environments.

Chapter 4 Advanced Feature Detection and Mapping Algorithm

4.1 Introduction

The major conclusion derived from the previous chapter as well as the previous research carried out in image-based AUV localisation emphasised that current feature detectors and descriptors are not able to work well in environments of an unstructured nature and with lack of significant features and comprising mostly repetitive features, especially in under-ice. This is a major problem faced in the AUV navigation and localisation community. Therefore, it is important to address the said problem with advanced feature detection and matching algorithm for underwater environments.

This chapter presents two novel combined algorithms as a substitute for the use of single algorithm aiming to detect more features in environments that lack distinguishable features. This can be done by combining different detectors and descriptors to detect more features in the environment and feature matching compared to using a single detector or descriptor. As discussed in the previous chapter, SURF performs well in almost all keypoint detection and matching compared to all other algorithms. In addition, Harris is capable of detecting the highest number of keypoints. The new advanced feature detection algorithms are named 'SURF-Harris', which combines Harris interest points with the SURF descriptor, and 'SURF-Harris-SURF', which combines Harris and SURF interest points with the SURF descriptor, using the most significant factor of each detector and descriptor to give better performance, especially in feature mapping. This is a novel contribution to the research on AUV localisation and mapping with sonar images.

4.2 Methodology

4.2.1 SURF-Harris (SUHa) Algorithm

This algorithm consists of a Harris detector and Spectra descriptor. The goal of Harris methodology is to find the direction of fast and lowest change of feature orientation, using covariance matrix of local directional derivatives. The directional derivative values are compared with the scoring factor to identify which features are corners, which are edges, and which are likely noise. Depending on the formulation of the algorithm, the Harris method can provide high rotational invariance, limited intensity invariance, and in some of the formulation of the algorithm, scale invariance is provided such as Harris-Laplace (Krig 2014b). The Harris corner detector is based on the local autocorrelation function of a signal, where the local autocorrelation function measures the local changes of the signal with patches shifted by a small amount in different directions. A discrete predecessor of the Harris detector was presented by Moravec (1980), where the discreteness refers to the shifting of the patches.

Given a shift $(\Delta x, \Delta y)$ and a point (x, y), the autocorrelation function is defined as shown in Eq. (4.1),

$$c(x,y) = \sum_{W} [I(x_i, y_i) - I(x_i + \Delta x, y_i + \Delta y)]^2 \tag{4.1}$$

where $I(.,.)$ denotes the image function and (x_i, y_i) are the points in the window W (Gaussian) centred on (x, y).

The shifted image is approximated by a Taylor expansion truncated to the first order terms as shown in Eq. (4.2),

$$I(x_i + \Delta x, y_i + \Delta y) \approx (x_i, y_i) + [I_x(x_i, y_i) I_y(x_i, y_i)] \begin{bmatrix} \Delta_x \\ \Delta_y \end{bmatrix} \tag{4.2}$$

where $I_x(.,.)$ and $I_y(.,.)$ denote the partial derivatives in x and y, respectively.

Substituting approximation Eq. (4.2) into Eq. (4.1) yields the following Eq. (4.3),

$$c(x,y) = \left[\Delta_x \Delta_y \right] C(x,y) \begin{bmatrix} \Delta_x \\ \Delta_y \end{bmatrix} \tag{4.3}$$

where matrix $C(x,y)$ captures the intensity structure of the local neighbourhood. Let λ1, λ2 be the eigenvalues of matrix $C(x,y)$. The eigenvalues form a rotationally invariant description. There are three cases to be considered:

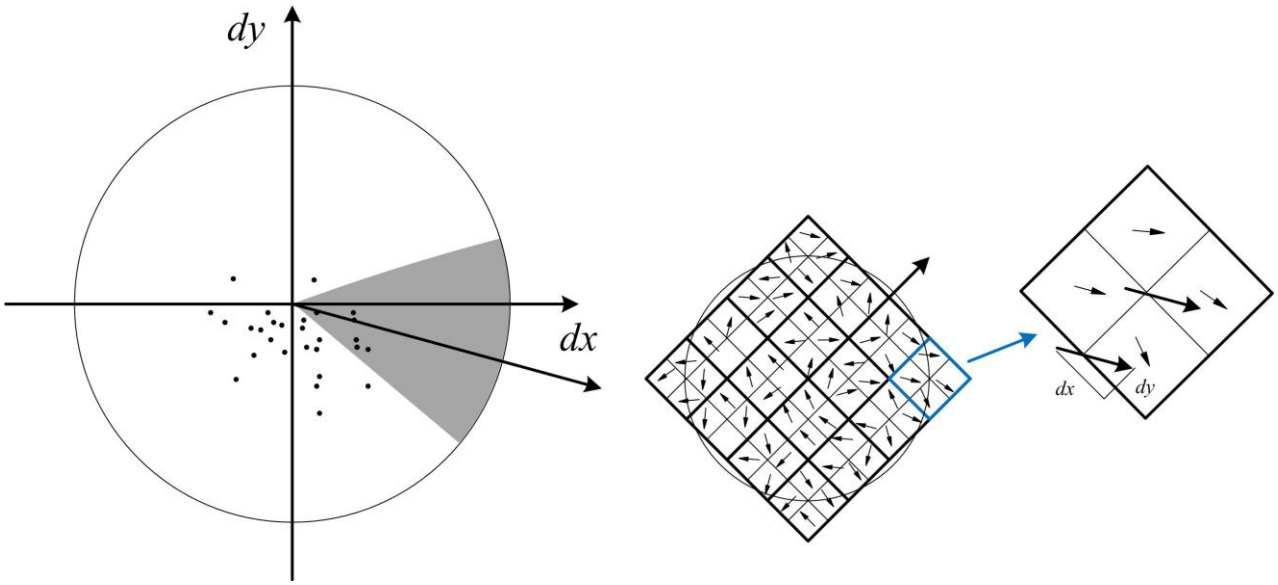
1. If both λ1, λ2 are small, so that the local autocorrelation function is flat (i.e., little change in $C(x,y)$in any direction), the windowed image region is of approximately constant intensity.

2. If one eigenvalue is high and the other low, so the local autocorrelation function is ridge shaped, then only local shifts in one direction (along the ridge) cause little change in $C(x,y)$ and significant change in the orthogonal direction; this indicates an edge.

3. If both eigenvalues are high, so the local autocorrelation function is sharply peaked, then shifts in any direction will result in a significant increase; this indicates a corner.

A detailed description of the Harris detector is to be found in Harris & Stephens (1988) and Derpanis (2004).

Moreover, this algorithm used the Spectra descriptor as a descriptor which is mostly involved in more intense computations. The descriptor is also based on the Haar wavelets and encodes the distribution of pixel intensity values in the neighbourhood of the detected features. Computation of the descriptor begins with the assignment of dominant orientation to make the descriptor rotation invarient. The sliding sector window of size π/3 is rotated around the interest point at intervals as shown in Figure 4-1. In this region, all are features are summed and this includes both the vertical and horizontal responses to produce a set of orientation vectors.



(a) (b) (c)

Figure 4-1. (a) The sliding sector window used in SURF to compute the dominant orientation of the Haar features to add rotational invariance to the SURF features. (b–c) The feature vector construction process, showing a grid containing a 4x4 region subdivided into 4x4 sub-regions and 2x2 (Bay et al. 2008).

To crete the descriptor vector for both algorithms, a rectangular grid of 4x4 regions is established surrounding the interest point (keypoints), and each region of this grid is split into 4x4 sub-regions. Within each sub-region, the Haar wavelet response is computed over 5x5 sample points. For each sub-region, the computation of Haar wavelet response in horizontal and vertival direction of vector V is represented by Eq. (4.4);

$$V = \left(\sum d_x, \sum d_y \, , \sum |d_x|, \sum |d_y| \right) \tag{4.4}$$

Derivatives in the x and y directions are taken in these sub-regions. The descriptor for the sub-region is the sum of the x derivatives over its four quardrants and similarly, y derivatives. For each sample point, Haar wavelet responses for two principal directionals are

computed. Each sub-region contributes to the descriptor with four components total vector length of 64 descriptor elements (Bay et al. 2006; Krig 2014a).

4.2.2 SURF-Harris-SURF Algorithm (SUHaSU)

This algorithm consists of Harris and Hessian matrix detectors with Spectra descriptor. As the Harris detector is discussed above, only the Hessian Matrix is explained below. The Hessian matrix, also referred to as the Determinant of Hessian (DoH) method, is also used in the SURF algorithm (Krig 2014c). It detects interest objects from a multi-scale image set where the determinant of the Hessian matrix is at a maxima and the Hessian matrix operator is calculated using the convolution of the second-order partial derivative of the Gaussian to yield a gradient maxima. The Hessian matrix $H(\mathbf{x},\sigma)$ of image *I,* given point of $\mathbf{x}$ = *(x, y)* and scale σ is defined in Eq. (4.5):

$$H(\mathbf{x},\sigma)=\begin{bmatrix} L_{xx}(\mathbf{x},\sigma) & L_{xy}(\mathbf{x},\sigma) \\ L_{xy}(\mathbf{x},\sigma) & L_{yy}(\mathbf{x},\sigma) \end{bmatrix} \tag{4.5}$$

where $L_{xx}(\mathbf{x},\sigma)$, $L_{xy}(\mathbf{x},\sigma)$,and $L_{yy}(\mathbf{x},\sigma)$ are the convolution of the Gaussian second order derivative with image *I* in point *x,* point *xy* and point *y* respectively. The approximated Hessian determinant symbolises the blob response of the image point at $\mathbf{x}$. Moreover, in the detection step the algorithm takes advantage of the use of Haar wavelets approximation of the blob detector based on the Hessian matrix determinant. Furthermore, this method uses integral images to calculate the Gausian partial derivatives very quickly. A detailed description of SURF is taken from Bay et al. (2006). In addition to this, calculated interest points are stored in matrix form as $P1$. Similarly, calculated interest points using the Harris method are stored in matrix as $P2$ as discussed in Section 4.2.1.

The total matrix P is obtained from Eq. (4.6),

$$P = P1 \cup P2 \tag{4.6}$$

For feature description, the above-mentioned Spectra descriptor is used as explained above. This descriptor produced an intensity contained distribution based on the *x* and *y* directions of 'Harr wavelet' responses within the neighbourhood area of an interest point which calculated in Eq.(4.6) (Bay et al. 2006).

4.3 Evaluation Method for Feature Detection and Mapping

For the validation of the proposed algorithms, this study was carried out with similar images to those used in the previous chapter. The study was conducted in the MATLAB 8.5.0 software environment where relevant tools are available from the computer vision system library. Similarly to the previous experiment in Chapter 3, there are four main steps carried out in image processing as shown in Figure 4-2.

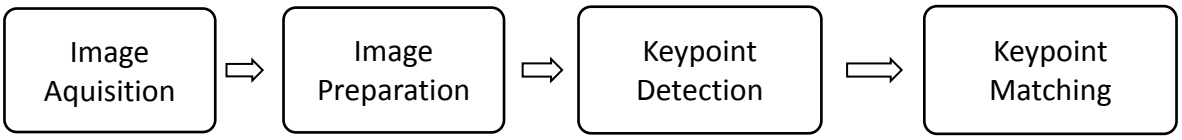
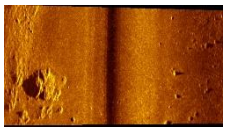
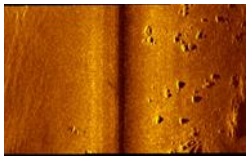
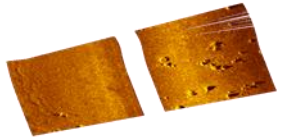


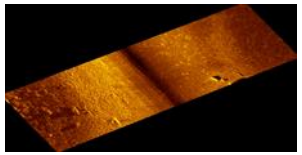
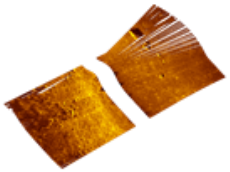
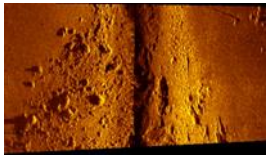
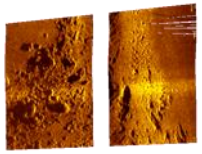
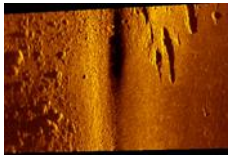
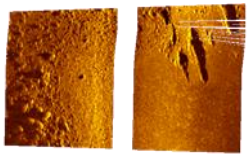
Figure 4-2. Flow chart in image processing

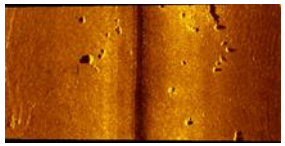
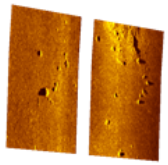
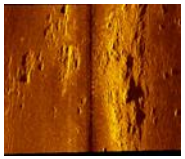
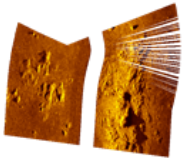
4.3.1 Image Acquisition

The images used in this study are given in Table 4-1. Each set consists of two images taken a day apart of the same area, similar to Table 3-1.

Table 4-1. Image data sets

Image Set	Images from Day 1	Images from day 2	Effects/Features
Data Set 1	 (Image 1)	 (Image 2)	Rotation Scale Differences in brightness Muddy surface with stones or sand Not in the same area
Data Set 2	 (Image 3)	 (Image 4)	Rotation Scale Muddy surface with stones or sand More overlapping area

Data Set 3	 (Image 5)	 (Image 6)	Rotation Scale Differences in brightness Muddy surface with stones or sand More overlapping area
Data Set 4	 (Image 7)	 (Image 8)	Scale Tilt Clutter Muddy surface with stones or sand More overlapping area
Data Set 5	 (Image 9)	 (Image 10)	Scale Muddy surface with stones or sand More overlapping area

Data Set 6	 (Image 11)	 (Image 12)	Scale Mud surface with stones or sand More overlapping area
Data Set 7	 (Image 13)	 (Image 14)	Scale Distortions Clutter Muddy surface with stones or sand More overlapping area

4.3.2 Image Pre-processing

As discussed in Chapter 3, all the images in Table 4-1 were subjected to resizing, filtering, contrast enhancement and morphological filtering.

4.3.3 Keypoint Detection

The pre-processed image is passed through the algorithms to detect keypoints. The algorithm selects regions of the image that have unique content, such as corners or blobs depending upon the detector using the algorithm. These detectors relied on gradient-based and intensity variation approaches to identify good local features. The important factor in the determination of the keypoint is to find key features that remain locally invariant so that it is possible to detect these key features even in the presence of rotation or scale change. Depending on the image, as well as the algorithm used, the number of keypoints detected can be as high as 1000 which may load the processor heavily in the next stage. The detected keypoints are stored in an m x n matrix form as per the example shown in Eq. (4.6) in the SURF-Harris-SURF algorithm and passed to the next stage for keypoint matching.

4.3.4 Keypoint Mapping

As mentioned in the previous sub-section 4.3.1, there are altogether seven data sets from Table 4-1, each having two images taken a day apart of the same area. Therefore, features of the two images in a set should match with each other. The matching process calculates the nearest neighbours between the identified features of the two images. If the best match's distance is less than a definite match of threshold value, the match is retained. Otherwise, the match is rejected. In this study, distinct match threshold value was set to 10.0 or 1.0 depending on their binary or non-binary feature vector. To get the best match between the features, an estimated geometric transformation has been used. M-estimator Sample Consensus (MSAC) algorithm is used to remove outliers in corresponding matched features which are a variant of the Random Sample Consequence (RANSAC) algorithm.

4.4 Results and Discussion

4.4.1 Keypoint Detection and Mapping

According to the previous discussion in sub-section 4.3.4, keypoints have been detected from all the pre-processed data set images using the 'SURF-Harris' algorithm and the 'SURF-Harris-SURF' algorithm as illustrated in Table 4-2. This section will not cover the details of processing images which have already been discussed in Chapter 3.

Table 4-2. Number of keypoints detected by SURF-Harris and SURF-Harris-SURF algorithms

Data Sets	Image Number	Day	SURF-Harris Algorithm	SURF-Harris-SURF Algorithm	
				Harris Detector	Hessian Detector
1	1	1	338	338	222
	2	2	307	307	169
2	3	1	488	488	214
	4	2	301	301	148
3	5	1	735	735	377
	6	2	618	618	193
4	7	1	315	315	649
	8	2	300	300	600
5	9	1	299	299	466
	10	2	295	295	438
6	11	1	321	321	204
	12	2	300	300	110
7	13	1	395	395	525
	14	2	299	299	460

As shown in Table 4-2, Harris and Hessian's ability to detect keypoints can be observed separately. These results proved that the Hessian detector produces the highest keypoint detection except for Data Sets 1, 2 and 3. The reason behind this is that Data Sets 4 to 7 have only scale differences. Furthermore, the Harris detector is capable of detecting more keypoints under rotation and difference in brightness. Moreover, these results proved that

the combination of detectors is better than a single detector using conventional feature detectors.

When two images are taken for the identification of keypoints, the maximum possible number of matches is decided by using the lowest number of detected keypoints of the two images. In every image set, the lower number is taken as the maximum possible number of matches, as shown in Table 4-3.

Table 4-3. Number of keypoints and mapping ratio for SURF-Harris and SURF-Harris-SURF algorithms

Data Sets	SURF-Harris Algorithm			SURF-Harris-SURF Algorithm		
	Keypoints	Keypoint Matching	Mapping Ratio %	Keypoints	Keypoint Matching	Mapping Ratio %
1	307	0	0.00	169	0	0
2	301	12	3.99	148	18	12.16
3	618	3	0.49	193	3	1.55
4	300	30	10.00	300	34	11.33
5	295	44	14.92	295	67	22.71
6	300	20	6.67	110	30	27.27
7	299	32	10.70	299	34	11.37

As depicted in Table 4-3, it is clear that the SURF-Harris-SURF algorithm works well compared to the SURF-Harris algorithm due to the combination of two detectors in a single algorithm. Furthermore, the SURF-Harris-SURF algorithm is able to maintain more than 10% of the matches and the matching percentage for Data Sets 1–2 and Data Sets 4–7. For example, even though the image Data Set 5 has the possibility of having 110 matches, the actual number of matches is 67, equal to 22.71 % matching. The corresponding matches are shown in Figure 4-3.

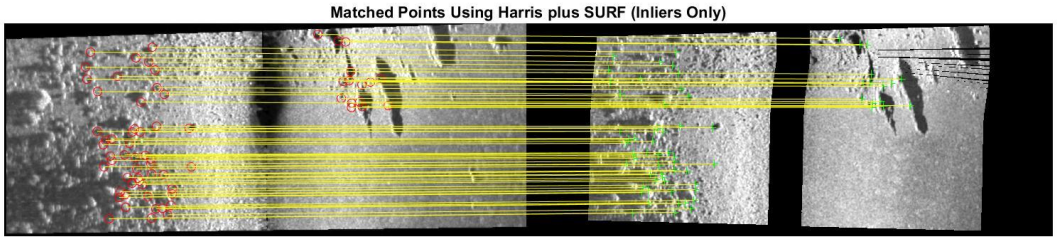


Figure 4-3. Keypoints matching using the 'SURF-Harris-SURF' algorithm for Data Set 5

Although the 'SURF-Harris' algorithm detected the highest number of keypoints, it failed to perform well in keypoint matching compared to the 'SURF-Harris-SURF' algorithm which used a combined detector. The major conclusion derived from these results is that the number of keypoint matches does not directly depend on the number of keypoints detection. The results analysis of Table 4-3 with Table 4-1 clearly shows the importance of feature distribution on the image in the matching process. As an example, after analysing Data Sets 5 and 6, both images have similar descriptions such as scale, overlapping and muddy surface. Nevertheless, Data Set 6 is able to detect more matches compared to Data Set 5 because of the feature distribution pattern of the image. Data Set 6 contains discrete features which are spread throughout the entire image, whereas Data Set 5 contains more features which are cluttered in some places.

4.5 Comparison of Results

This section discusses the performance comparison of the 'SURF-Harris' and 'SURF-Harris-SURF' algorithms with a single feature detector and descriptor such as SURF, Harris, FAST, BRISK or FREAK algorithms. When it comes to the feature matching of each algorithm, the detected keypoints in Day 1 should match with detected keypoints in Day 2 as illustrated in Figure 4-4.

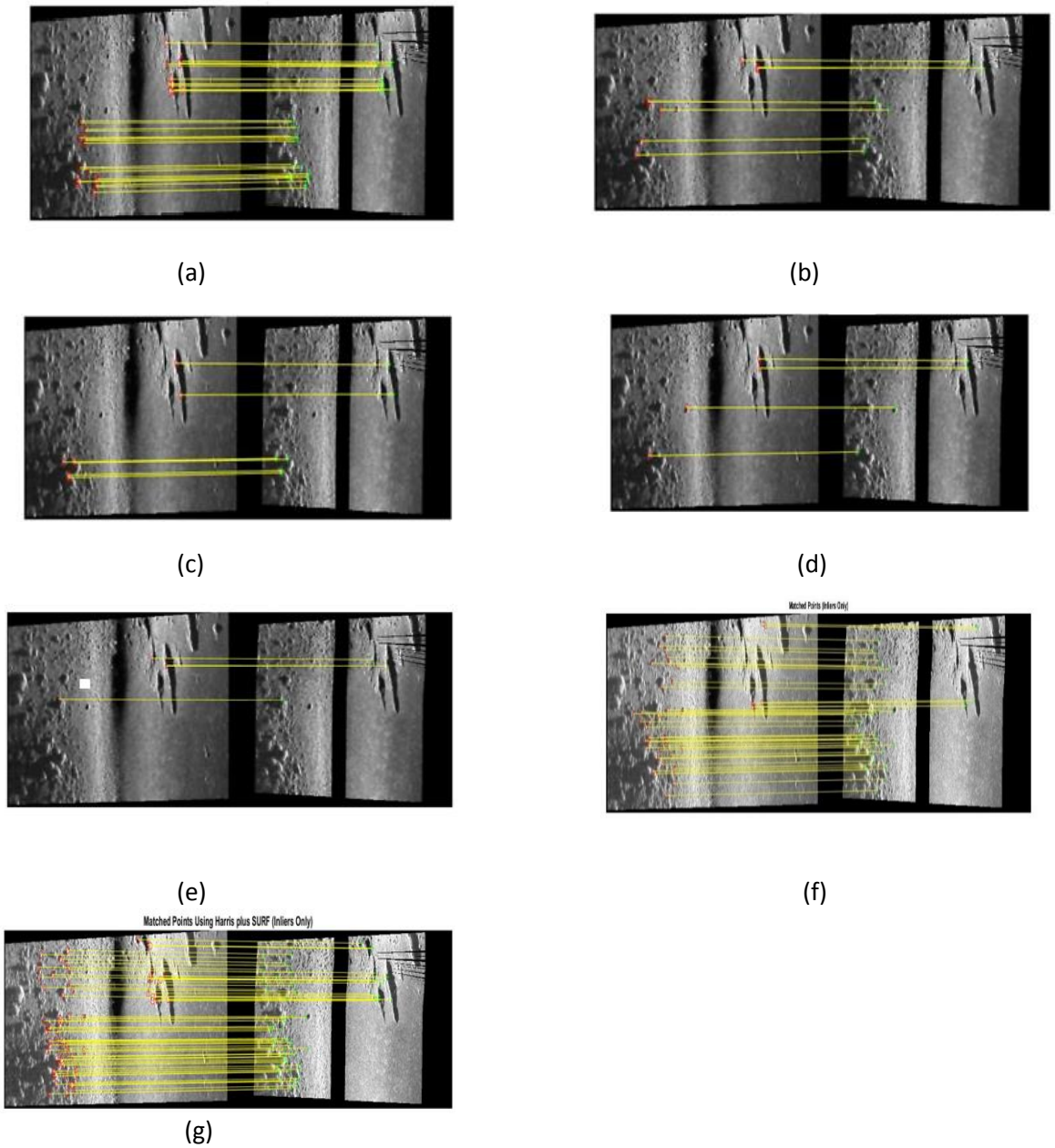


Figure 4-4. Keypoints matching using Data Set 5 with (a) SURF algorithm; (b) FAST algorithm; (c) Harris algorithm; (d) BRISK algorithm; (e) FREAK algorithm; (f) SURF-Harris algorithm; (g) SURF-Harris-SURF algorithm

The number of keypoint matches with all the data sets in the 'SURF-Harris' algorithm and 'SURF-Harris-SURF' algorithm with other conventional detectors and descriptors are illustrated in Figure 4-5. The major identification that can be derived from Figure 4-5 is that there is a higher number of keypoint matches obtained from 'SURF-Harris' and 'SURF-Harris-SURF' algorithms compared to the conventional feature detector and descriptor algorithms. As shown in Figure 4.5, Data Set 1 does not have a single matching feature and this is confirmed by the naked eye from the data set. Similarly, amidst rotations and scaling, the image Data Sets 2–7 can be seen as matching sets for the naked eye with clearly distinguishable features. This is confirmed by the results having a varied number of matches produced by each method. The results in Figure 4-5 further reveal that the 'SURF-Harris' and 'SURF-Harris-SURF' algorithms perform well in all the cases producing the highest number of matches.

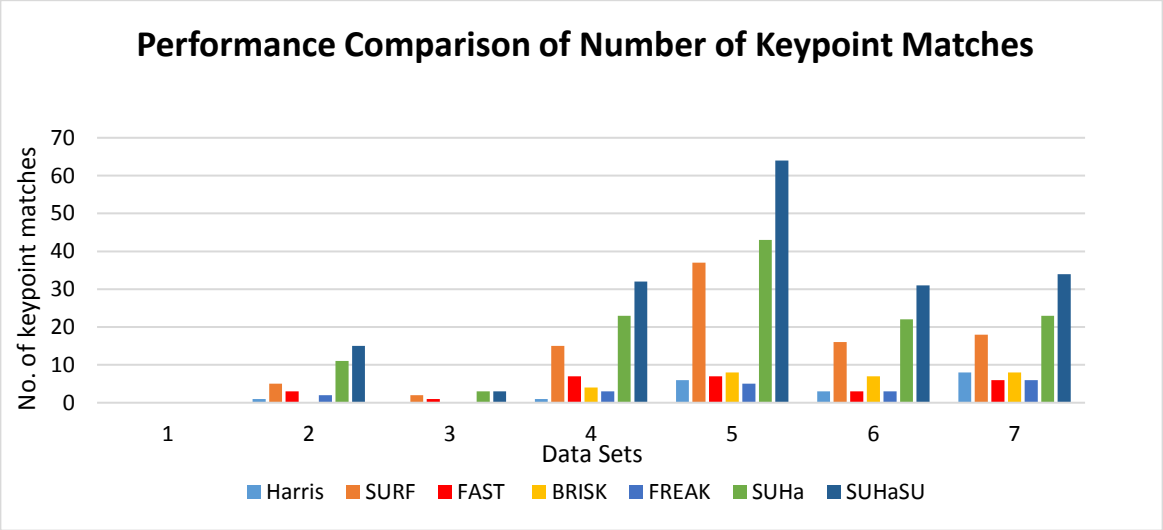


Figure 4-5. Number of matches produced by each algorithm for image sets

The major conclusion that can be derived from the results is that the new algorithms outperform all the other methods in keypoint matching even in the presence of scaling, rotation and differences in image intensity due to the combination of each detector and descriptor based on their own capability. This is a novel contribution in underwater environments with sonar images. Moreover, these results proved that new algorithms perform well compared to the conventional feature detectors and descriptors such as SURF, Harris, FAST, BRISK and FREAK. Even if there is a large number of features in a cluttered environment, it produces less matching compared to the features that are distributed.

Therefore, another conclusion which can be derived from the results is that the feature detection and matching algorithms performed well in discrete environments.

Chapter 5 **Summary, Conclusion and Future Work**

This chapter provides an overall summary of the thesis and brings together the findings of the individual chapters. It also concludes the findings and outcomes, and discusses the implications of the findings, the limitations, and provides recommendations for further research.

5.1 Summary

This thesis investigated the current problems in Autonomous Underwater Vehicle (AUV) navigation in under-ice environments using side-scan sonars and focus to develop new feature extraction algorithms to work in side-scan sonar images in such environments. In this chapter, an overall evaluation is made of the results and findings, and their contributions to the research field. Limitations of the study are also discussed and used to provide guidance for future research to increase the understanding of AUV navigation in underwater or under-ice environments by using mature technology side-scan sonar images.

A review of recent advances in under-ice localisation and navigation has been performed in terms of various sensors and algorithms, along with their pros and cons. The key challenge in AUV navigation is to identify the reliable features in low-contrast environments with a lack of features. Extracting good features under-ice is the most challenging task. This requires advanced feature detection methods for sonar images to work in those environments. Moreover, compared to making improvements to individual detection methods, exploration of their possible combinations would yield better results.

Therefore, this thesis has presented a performance comparison of SURF, Harris, FAST, BRISK and FREAK algorithms adding filtering effect in keypoint detection and matching of sonar images. The reported work shows that each detector and descriptor is capable of handling images from various disciplines such as rotation, scale and brightness. In addition to that, these detectors and descriptors work with sonar images in four different feature categories such as fewer features with clutter, fewer features with sparse arrangement, a high number of features with clutter and a high number of features with sparse arrangement.

These results prove that SURF performs well in all the data sets where BRISK is failing to work images with noise and rotation. However, FAST managed to work with all the data sets after SURF. Harris is suitable for feature detection as compared to the mapping process. FREAK performed well in rotation and scale images and failed to work with brightness environment of images due to changes in the image intensity for retina sampling pattern. Furthermore, the number of feature detectors and descriptors such as SURF, Harris, FAST, BRISK and FREAK algorithms work well in environments with features. However, these detectors and descriptors are not able to work effectively in environments that lack features. Therefore, this thesis has further studied to address this problem by developing new advanced algorithms named 'SURF-Harris' algorithm, which combined Harris interest points with the SURF descriptor, and 'SURF-Harris-SURF' algorithm, which combined Harris and SURF interest points with the SURF descriptor, using the most significant factor of each detector and descriptor to give better performance, especially in feature mapping. Compared to conventional feature detectors and descriptors, these two algorithms are able to work well in underwater environments that lack significant textures and distinguishing patterns.

5.2 Conclusion

The major contribution of this study is the introduction of two new feature detection algorithms called ‘SURF-Harris’ and ‘SURF-Harris-SURF’ to the field of computer vision, especially image processing. The ‘SURF-Harris-SURF’ algorithm outperforms all the other methods in feature matching with filtering, even in the presence of scaling, rotation, and differences in image intensity. Moreover, the results documented in this thesis have proved that the new algorithms perform well in comparison to the conventional feature detectors and descriptors such as SURF, Harris, FAST, BRISK and FREAK. These results have shown that as a substitute for using a single detector, a combination of several detectors detects more features in such environments. This is the novel contribution of this research.

Moreover, the SURF algorithm works well in all the disciplines with a higher percentage of matching even though it produces fewer keypoints, thus demonstrating its robustness among all the conventional detectors and descriptors. Furthermore, it has been shown that the number of detected keypoints reduced while using filtering effect. Nevertheless, it has been proved that the number of true key matches increased with filtering. It is also concluded that the distribution of features within the image has a significant impact on the matching process. Even if there is large number of features in a cluttered environment, it produces less matching compared to features that are distributed. Therefore, another conclusion that can be derived from the results is that the feature detection and matching algorithms performed well in sparse environments.

5.3 Future Work

Although these new algorithms work in underwater environments, they are not able to give 100% efficiency, especially in environments with lack of features, and this is an area to be further explored. Therefore, it is important to address this problem with the new advanced algorithms for feature detection and extraction, especially for creating sonar images. There are various types of detector methods available such as Laplacian of Gaussian (LOG), Harris and Stephens corner detection, Difference of Gaussian (DOG), Determinant of Hessian (DoH), Salient Regions, SUSAN, and Morphological interest points. Each detector has its own advantages and disadvantages when working in different environments. Therefore, the detector is going to be selected according to its capability to work with low contrast and environments with a lack of features. Similarly, according to the literature, there are various descriptors available such as Local Binary descriptors, Spectra descriptors and basic space descriptors. Out of these, the most suitable descriptor is going to be developed further to work with sonar images and build new descriptors according to the described environment. This can be a new research area: to develop advanced feature detection algorithms to work with low contrast sonar images without using existing detectors or descriptors. As this is an area that has not been researched to date, it is a good area for further research. As with present AUV operations, navigation is a crucial element for long-range under-ice missions.

Noise suppression and image smoothing in the process of enhancement are very important factors in image processing in sonar images. Therefore, the development of new techniques for optimising the enhancement process of side-scan sonar images is suggested as a new area of research development in the future, as is the development of a better computation methodology for calculating threshold value for side-scan sonar images, especially with non-uniform illuminations. Furthermore, the developed algorithms of this thesis can be used for AUV navigation with a combination of SLAM and TRN. A new algorithm or an extended version of the available algorithm for SLAM could be developed in order to reduce computation time. The further development of a path-searching algorithm would reduce the current estimated error in pre-survey area of the vehicle location. Another possibility to be developed is map building using side-scan sonar. This method is a similar

method to that which has been carried out for the translation and rotation of free floating icebergs side-scan sonars instead of multi-beam sonars.

Author's publications

BANDARA, D., LEONG, Z., NGUYEN, H., JAYASINGHE, S. & FORREST, A. L. Technologies for under-ice AUV navigation. Autonomous Underwater Vehicles (AUV), 2016 IEEE/OES, 2016. IEEE, 108-114.

References

ALAHI, A., ORTIZ, R. & VANDERGHEYNST, P. Freak: Fast retina keypoint. Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, 2012. Ieee, 510-517.

BANDARA, D., LEONG, Z., NGUYEN, H., JAYASINGHE, S. & FORREST, A. L. Technologies for under-ice AUV navigation. Autonomous Underwater Vehicles (AUV), 2016 IEEE/OES, 2016. IEEE, 108-114.

BAY, H., ESS, A., TUYTELAARS, T. & VAN GOOL, L. 2008. Speeded-up robust features (SURF). *Computer vision and image understanding,* 110**,** 346-359.

BAY, H., TUYTELAARS, T. & GOOL, L. 2006. SURF: Speeded Up Robust Features. *In:* LEONARDIS, A., BISCHOF, H. & PINZ, A. (eds.) *Computer Vision – ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I.* Berlin, Heidelberg: Springer Berlin Heidelberg.

BLACKINTON, J., HUSSONG, D. & KOSALOS, J. First results from a combination side-scan sonar and seafloor mapping system (SeaMARC II). Offshore Technology Conference, 1983. Offshore Technology Conference.

BLONDEL, P. 2010. *The handbook of sidescan sonar*, Springer Science & Business Media.

BURGUERA, A., BONIN-FONT, F. & OLIVER, G. 2015. Trajectory-Based Visual Localization in Underwater Surveying Missions. *Sensors,* 15**,** 1708-1735.

CALONDER, M., LEPETIT, V., STRECHA, C. & FUA, P. 2010. Brief: Binary robust independent elementary features. *Computer Vision–ECCV 2010***,** 778-792.

CARESS, D. W., CLAGUE, D. A., PADUAN, J. B., MARTIN, J. F., DREYER, B. M., CHADWICK JR, W. W., DENNY, A. & KELLEY, D. S. 2012. Repeat bathymetric surveys at 1-metre resolution of lava flows erupted at Axial Seamount in April 2011. *Nature Geoscience,* 5**,** 483-488.

CHEN, WANG, S., MCDONALD ‐ MAIER, K. & HU, H. 2013. Towards autonomous localization and mapping of AUVs: a survey. *International Journal of Intelligent Unmanned Systems,* 1**,** 97-120.

CHEN, L. & HU, H. 2011. Towards localization and mapping of autonomous underwater vehicles: A survey.

CHEN, P. Y., LI, Y., SU, Y. M., CHEN, X. L. & JIANG, Y. Q. 2015. Review of AUV Underwater Terrain Matching Navigation. *Journal of Navigation,* 68**,** 1155-1172.

CLAUS, B. & BACHMAYER, R. Towards online terrain aided navigation of underwater gliders. 2014 IEEE/OES Autonomous Underwater Vehicles (AUV), 6-9 Oct. 2014 2014. 1-5.

DEFFENBAUGH, M., SCHMIDT, H. & BELLINGHAM, J. G. Acoustic navigation for Arctic under-ice AUV missions. OCEANS '93. Engineering in Harmony with Ocean. Proceedings, 18-21 Oct 1993 1993. I204-I209 vol.1.

DERPANIS, K. G. 2004. The harris corner detector. *York University*.

DI MASSA, D. E. & STEWART JR, W. Terrain-relative navigation for autonomous underwater vehicles. OCEANS'97. MTS/IEEE Conference Proceedings, 1997. IEEE, 541-546.

FAIRFIELD, N., KANTOR, G. & WETTERGREEN, D. 2007. Real ‐ Time SLAM with Octree Evidence Grids for Exploration in Underwater Tunnels. *Journal of Field Robotics,* 24**,** 03-21.

FORREST, A., HAMILTON, A., SCHMIDT, V., LAVAL, B., MUELLER, D., CRAWFORD, A., BRUCKER, S. & HAMILTON, T. Digital terrain mapping of Petermann Ice Island fragments in the Canadian high arctic. 21st IAHR International Symposium on Ice, 2012. 1-12.

FORREST, A. L., LAVAL, B., DOBLE, M. J., YEO, R. & MAGNUSSON, E. AUV measurements of under-ice thermal structure. OCEANS 2008, 15-18 Sept. 2008 2008. 1-10.

FRANCOIS, R. E., NODLAND, W. E. & LABORATORY, U. O. W. A. P. 1972. *Unmanned Arctic Research Submersible (UARS) System Development and Test Report*, Applied Physics Laboratory, University of Washington.

FRAUNDORFER, F. & SCARAMUZZA, D. 2012. Visual Odometry : Part II: Matching, Robustness, Optimization, and Applications. *IEEE Robotics & Automation Magazine,* 19**,** 78-90.

FREITAG, L., GRUND, M., VON ALT, C., STOKEY, R. & AUSTIN, T. 2005. A shallow water acoustic network for mine countermeasures operations with autonomous underwater vehicles. *Underwater Defense Technology (UDT)***,** 1-6.

GOLDEN, J. P. 1980. Terrain contour matching (TERCOM): A cruise missile guidance aid. *Image processing for missile guidance,* 238**,** 10-18.

GONZALEZ, R. C. & WOODS, R. E. 2002. *Digital image processing,* United States of America, PEARSON.

GREWAL, M. S., WEILL, L. R. & ANDREWS, A. P. 2007. *Global positioning systems, inertial navigation, and integration*, John Wiley & Sons.

GUTH, F., SILVEIRA, L., BOTELHO, S., DREWS, P. & BALLESTER, P. Underwater SLAM: Challenges, state of the art, algorithms and a new biologically-inspired approach. Biomedical Robotics and Biomechatronics (2014 5th IEEE RAS & EMBS International Conference on, 12-15 Aug. 2014 2014. 981-986.

HARRIS, C. & STEPHENS, M. A combined corner and edge detector. 1988. Citeseer.

HASSABALLAH, M., ABDELMGEID, A. A. & ALSHAZLY, H. A. 2016. Image Features Detection, Description and Matching. *Image Feature Detectors and Descriptors.* Springer.

HE, B., LIANG, Y., FENG, X., NIAN, R., YAN, T., LI, M. & ZHANG, S. 2012. AUV SLAM and experiments using a mechanical scanning forward-looking sonar. *Sensors,* 12**,** 9386-9410.

HILDEBRANDT, M., ALBIEZ, J., FRITSCHE, M., HILLJEGERDES, J., KLOSS, P., WIRTZ, M. & KIRCHNER, F. Design of an autonomous under-ice exploration system. Oceans - San Diego, 2013, 23-27 Sept. 2013 2013. 1-6.

HOU, S., PENG, S., YAN, Z. & ZHANG, W. Research on the error model of ins/dvl system for autonomous underwater vehicle. Automation and Logistics, 2008. ICAL 2008. IEEE International Conference on, 2008. IEEE, 201-206.

JACOFF, A., SAIDI, K., LOEWENFELDT, R. V. & KOIBUCHI, Y. Development of standard test methods for evaluation of ROV/AUV performance for emergency response applications. OCEANS 2015 - MTS/IEEE Washington, 19-22 Oct. 2015 2015. 1-10.

JAKUBA, M. V., ROMAN, C. N., SINGH, H., MURPHY, C., KUNZ, C., WILLIS, C., SATO, T. & SOHN, R. A. 2008. Long‐baseline acoustic navigation for under‐ice autonomous underwater vehicle operations. *Journal of Field Robotics,* 25**,** 861-879.

KAMINSKI, C., CREES, T., FERGUSON, J., FORREST, A., WILLIAMS, J., HOPKIN, D. & HEARD, G. 12 days under ice – an historic AUV deployment in the Canadian High Arctic. 2010 IEEE/OES Autonomous Underwater Vehicles, 1-3 Sept. 2010 2010. 1-11.

KIMBALL, P. & ROCK, S. 2011. Sonar-based iceberg-relative navigation for autonomous underwater vehicles. *Deep-Sea Research Part Ii-Topical Studies in Oceanography,* 58**,** 1301-1310.

KIMBALL, P. W. & ROCK, S. M. 2015. Mapping of Translating, Rotating Icebergs With an Autonomous Underwater Vehicle. *Ieee Journal of Oceanic Engineering,* 40**,** 196-208.

KING, P., ANSTEY, B. & VARDY, A. Comparison of feature detection techniques for auv navigation along a trained route. Oceans-San Diego, 2013, 2013. IEEE, 1-8.

KING, P., ANSTEY, B. & VARDY, A. Preliminary field trials of autonomous path following. Autonomous Underwater Vehicles (AUV), 2014 IEEE/OES, 6-9 Oct. 2014 2014. 1-7.

KING, P., ANSTEY, B. & VARDY, A. 2017. SONAR IMAGE REGISTRATION FOR LOCALIZATION OF AN UNDERWATER VEHICLE. *Journal of Ocean Technology,* 12.

KING, P., VARDY, A., VANDRISH, P. & ANSTEY, B. Real-time side scan image generation and registration framework for AUV route following. Autonomous Underwater Vehicles (AUV), 2012 IEEE/OES, 24-27 Sept. 2012 2012. 1-6.

KINSEY, EUSTICE & WHITCOMB. A survey of underwater vehicle navigation: Recent advances and new challenges. IFAC Conference of Manoeuvering and Control of Marine Craft, 2006.

KRIG, S. 2014a. *Computer vision metrics*, Springer.

KRIG, S. 2014b. Interest point detector and feature descriptor survey. *Computer Vision Metrics.* Springer.

KRIG, S. 2014c. Interest Point Detector and Feature Descriptor Survey. *Computer Vision Metrics: Survey, Taxonomy, and Analysis.* Berkeley, CA: Apress

KUKULYA, A., PLUEDDEMANN, A., AUSTIN, T., STOKEY, R., PURCELL, M., ALLEN, B., LITTLEFIELD, R., FREITAG, L., KOSKI, P. & GALLIMORE, E. Under-ice operations with a REMUS-100 AUV in the Arctic. Autonomous Underwater Vehicles (AUV), 2010 IEEE/OES, 2010. IEEE, 1-8.

LEUTENEGGER, S., CHLI, M. & SIEGWART, R. Y. BRISK: Binary robust invariant scalable keypoints. Computer Vision (ICCV), 2011 IEEE International Conference on, 2011. IEEE, 2548-2555.

LI, Y., MA, T., CHEN, P., JIANG, Y., WANG, R. & ZHANG, Q. 2017. Autonomous underwater vehicle optimal path planning method for seabed terrain matching navigation. *Ocean Engineering,* 133**,** 107-115.

LIN, Z. & WEI, G. The experimental study on GPS/INS/DVL integration for AUV. Position Location and Navigation Symposium, 2004. PLANS 2004, 26-29 April 2004 2004. 337-340.

LOWE, D. G. Object recognition from local scale-invariant features. Computer vision, 1999. The proceedings of the seventh IEEE international conference on, 1999. Ieee, 1150-1157.

MARQUES, O. 2011. *Practical image and video processing using MATLAB*, John Wiley & Sons.

MATLAB. 2017. Available: https://au.mathworks.com/help/vision/local-feature-extraction.html [Accessed 10/01/2017 2017].

MCCARTHY, E. M. & SABOL, B. Acoustic characterization of submerged aquatic vegetation: military and environmental monitoring applications. OCEANS 2000 MTS/IEEE Conference and Exhibition, 2000. IEEE, 1957-1961.

MCEWEN, R., THOMAS, H., WEBER, D. & PSOTA, F. 2005. Performance of an AUV navigation system at Arctic latitudes. *Oceanic Engineering, IEEE Journal of,* 30**,** 443-454.

MCFARLAND, C. J., JAKUBA, M. V., SUMAN, S., KINSEY, J. C. & WHITCOMB, L. L. Toward ice-relative navigation of underwater robotic vehicles under moving sea ice: Experimental evaluation in the Arctic sea. Robotics and Automation (ICRA), 2015 IEEE International Conference on, 26-30 May 2015 2015. 1527-1534.

MEDAGODA, L., WILLIAMS, S. B., PIZARRO, O., KINSEY, J. C. & JAKUBA, M. V. 2016. Mid-water current aided localization for autonomous underwater vehicles. *Autonomous Robots,* 40**,** 1207-1227.

MEDUNA, D. K., ROCK, S. M. & MCEWEN, R. S. Closed-loop terrain relative navigation for AUVs with non-inertial grade navigation sensors. 2010 IEEE/OES Autonomous Underwater Vehicles, 1-3 Sept. 2010 2010. 1-8.

MILLER, P. A., FARRELL, J. A., ZHAO, Y. & DJAPIC, V. 2010. Autonomous Underwater Vehicle Navigation. *IEEE Journal of Oceanic Engineering,* 35**,** 663-678.

MORAVEC, H. P. 1980. Obstacle avoidance and navigation in the real world by a seeing robot rover. DTIC Document.

NGUYEN, H. G., FABLET, R., EHRHOLD, A. & BOUCHER, J. M. 2012. Keypoint-Based Analysis of Sonar Images: Application to Seabed Recognition. *IEEE Transactions on Geoscience and Remote Sensing,* 50**,** 1171-1184.

PADIAL, J., DEKTOR, S. & ROCK, S. M. Correlation of imaging sonar acoustic shadows and bathymetry for ROV terrain-relative localization. OCEANS 2014-TAIPEI, 2014. IEEE, 1-10.

PADIAL, J., DEKTOR, S. G. & ROCK, S. M. 2013. Correlation of sidescan sonar acoustic shadows and bathymetry for terrain-relative navigation. *Unmanned Untethered Submersible Technology*.

PAULL, L., SAEEDI, S., SETO, M. & LI, H. 2014. AUV Navigation and Localization: A Review. *Oceanic Engineering, IEEE Journal of,* 39**,** 131-149.

PINTO, M., FERREIRA, B., MATOS, A. & CRUZ, N. Using side scan sonar for relative navigation. Industrial Electronics, 2009. IECON'09. 35th Annual Conference of IEEE, 2009. IEEE, 2126-2141.

RICHMOND, K., GULATI, S., FLESHER, C., HOGAN, B. P. & STONE, W. C. Navigation, control, and recovery of the ENDURANCE under-ice hovering AUV. International Symposium on Unmanned Untethered Submersible Technology UUST, 2009. Citeseer.

RIGBY, P., PIZARRO, O. & WILLIAMS, S. B. Towards Geo-Referenced AUV Navigation Through Fusion of USBL and DVL Measurements. OCEANS 2006, 18-21 Sept. 2006 2006. 1-6.

ROMAN, C. & SINGH, H. Improved vehicle based multibeam bathymetry using sub-maps and SLAM. Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on, 2-6 Aug. 2005 2005. 3662-3669.

ROSTEN, E. & DRUMMOND, T. 2006. Machine Learning for High-Speed Corner Detection. *In:* LEONARDIS, A., BISCHOF, H. & PINZ, A. (eds.) *Computer Vision – ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I.* Berlin, Heidelberg: Springer Berlin Heidelberg.

SCHMIDT, A. & KRAFT, M. 2015. The Impact of the Image Feature Detector and Descriptor Choice on Visual SLAM Accuracy. *In:* CHORAŚ, S. R. (ed.) *Image Processing & Communications Challenges 6.* Cham: Springer International Publishing.

SCHMIDT, A., KRAFT, M., FULARZ, M. & DOMAGAŁA, Z. 2013. Comparative assessment of point feature detectors in the context of robot navigation. *Journal of Automation Mobile Robotics and Intelligent Systems,* 7**,** 11--20.

SCHOFIELD, O., DUCKLOW, H. W., MARTINSON, D. G., MEREDITH, M. P., MOLINE, M. A. & FRASER, W. R. 2010. How do polar marine ecosystems respond to rapid climate change? *Science,* 328**,** 1520-1523.

SINGH, H., YOERGER, D., BACHMAYER, R., BRADLEY, A. & STEWART, W. K. Sonar mapping with the autonomous benthic explorer (ABE). INTERNATIONAL SYMPOSIUM ON UNMANNED UNTETHERED SUBMERSIBLE TECHNOLOGY, 1995. UNIVERSITY OF NEW HAMPSHIRE-MARINE SYSTEMS, 367-375.

SMITH, S. 1992. A New Class of Corner Finder. *In:* HOGG, D. & BOYLE, R. (eds.) *BMVC92: Proceedings of the British Machine Vision Conference, organised by the British Machine Vision Association 22–24 September 1992 Leeds.* London: Springer London.

SONG, Y., HE, B., ZHANG, L. & YAN, T. Side-scan sonar image registration based on modified phase correlation for AUV navigation. OCEANS 2016 - Shanghai, 10-13 April 2016 2016. 1-4.

SPEARS, A., HOWARD, A., WEST, M. & COLLINS, T. Sonar and video fusion for vehicle trajectory estimation in under-ice environments. OCEANS 2015 - MTS/IEEE Washington, 19-22 Oct. 2015 2015. 1-8.

STALDER, S., BLEULER, H. & URA, T. Terrain-based navigation for underwater vehicles using side scan sonar images. OCEANS 2008, 15-18 Sept. 2008 2008. 1-3.

STONE AEROSPACE/PSC, I. 2016. *Environmentally Non-Disturbing Under-ice Robotic ANtarctiC Explorer* [Online]. Available: http://stoneaerospace.com/endurance/ [Accessed 15.06.2016 2016].

VANDRISH, P., VARDY, A., WALKER, D. & DOBRE, O. A. Side-scan sonar image registration for AUV navigation. 2011 IEEE Symposium on Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies, 5-8 April 2011 2011. 1-7.

VESTGÅRD, K. 1985. Underwater navigation and positioning systems. *Arctic Underwater Operations.* Springer.

WEBSTER, S. E., FREITAG, L. E., LEE, C. M. & GOBAT, J. I. Towards real-time under-ice acoustic navigation at mesoscale ranges. Robotics and Automation (ICRA), 2015 IEEE International Conference on, 26-30 May 2015 2015. 537-544.

WIENER, N. 1949. *Extrapolation, interpolation, and smoothing of stationary time series*, MIT press Cambridge, MA.

WILLIAMS, D., KRISHNALAL, G. & RAJ, V. J. 2016. Fraudulent Image Recognition Using Stable Inherent Feature. *Advances in Signal Processing and Intelligent Recognition Systems.* Springer.

WILLIAMS, S. B., PIZARRO, O. R., JAKUBA, M. V., JOHNSON, C. R., BARRETT, N. S., BABCOCK, R. C., KENDRICK, G. A., STEINBERG, P. D., HEYWARD, A. J., DOHERTY, P. J., MAHON, I., JOHNSON-ROBERSON, M., STEINBERG, D. & FRIEDMAN, A. 2012. Monitoring of Benthic Reference Sites: Using an Autonomous Underwater Vehicle. *IEEE Robotics & Automation Magazine,* 19**,** 73-84.

WOOCK, P. & FREY, C. Deep-sea AUV navigation using side-scan sonar images and SLAM. OCEANS 2010 IEEE-Sydney, 2010. IEEE, 1-8.

ZERR, B., MAILFERT, G., BERTHOLOM, A. & AYREAULT, H. Sidescan sonar image processing for AUV navigation. Europe Oceans 2005, 20-23 June 2005 2005. 124-130 Vol. 1.